

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:38:11 ; Search time 17.7526 Seconds
(without alignments)
444.049 Million cell updates/sec

Title: US-09-768-826-47_COPY_24_105
Perfect score: 427
Sequence: 1 EDNDEFFMDFLQTLVGTPE.....VQVLGSQDAGTDRKDDDK 82

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	67.9	94	2 S17449	probable ligand-bi
2	85.5	20.0	96	2 A36581	polychlorinated bi
3	79.5	18.6	91	2 J50036	Clara cell 10K pro
4	76.5	17.9	95	2 S68231	FHG22 protein prec
5	74.5	17.4	91	1 UGRB	uteroglobin precu
6	69.5	16.3	96	1 UGMS	uteroglobin precu
7	69.5	16.3	2458	2 T17420	probable polyketid
8	69	16.2	540	2 G86790	fibronectin-bindin
9	68.5	16.0	2100	2 T38128	t123.15 protein -
10	67.5	15.8	113	2 JC2026	cell specific 10K
11	67.5	15.8	1953	2 G3244	BN11 protein - yea
12	67	15.7	270	2 S30947	hypothetical prote
13	67	15.7	423	2 UGRBL	mannose-6-phosphat
14	66.5	15.6	91	1 UGRBL	uteroglobin precu
15	66.5	15.6	358	2 A3580	iron(III)-binding
16	66	15.5	192	2 S70285	outer surface prot
17	65.5	15.3	362	2 A71243	probable maltose/m
18	65.5	15.3	609	2 T38656	probable RNA-bind
19	65.5	15.3	1687	2 T30244	phosphodiesterase
20	65.5	15.3	1706	2 T30175	exoribonuclease, v
21	65.5	15.3	1719	2 T30174	exoribonuclease, v
22	65	15.2	209	2 S73690	MG331 homolog P01
23	65	15.2	255	2 B64175	repressor mode hom
24	65	15.2	312	2 G71146	hypothetical prote
25	65	15.2	842	2 E96641	hypothetical prote
26	64.5	15.1	339	2 E30436	coA-ligase / coenz
27	64.5	15.1	403	2 B98908	hypothetical prote
28	64.5	15.1	454	1 XNRTY	tyrosine transamin
29	64	15.0	342	1 C69395	H+-transporting AT

30 64 15.0 627 2 H96951 fusion, PTS system
31 64 15.0 6713 2 B89921 hypothetical prote
32 63.5 14.9 1418 2 T15232 hypothetical prote
33 63.5 14.9 1939 2 T18372 repeat organellar
34 63 14.8 176 2 B64549 conserved hypotet
35 63 14.8 454 2 S69017 probable membrane
36 63 14.8 3187 2 JC5837 364K Golgi complex
37 62.5 14.6 258 1 OYECHF imidazoleglycerol-
38 62.5 14.6 258 2 B90982 imidazoleglycerol-
39 62.5 14.6 258 2 H85827 imidazoleglycerol-
40 62.5 14.6 362 2 A75208 sugar transport AT
41 62.5 14.6 884 2 D96730 unknown protein F5
42 62 14.5 210 2 S69923 outer surface prot
43 62 14.5 744 2 A54901 long-chain-fatty-a
44 61.5 14.4 410 2 E75190 probable phosphono
45 61.5 14.4 1040 2 D81379 transmembrane effl

ALIGNMENTS

RESULT 1
S17449
probable ligand-binding protein RYD5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995-#sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S17449
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.
EMBO J. 10, 2813-2819, 1991
A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac
A:Reference number: S17447; MUID:92007724; PMID:1915264
A:Accession: S17449
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-94 <DEA>
A:Cross-references: EMBL:X60661; NID:957735; PIDN:CAA43068.1; PID:957736

Query Match 67.9% Score 290; DB 2; Length 94;
Best Local Similarity 80.0% Pred. No. 1.1e-24;
Matches 56; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDNDEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLOPMKAEVLVK 60
Db 23 EDDNEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLOPMKAEVLVK 82
Qy 61 LLVQVLGSQD 70
Db 83 LLVQVLDAQE 92

RESULT 2
A36581
Polychlorinated biphenyl-binding protein precursor - rat
N:Alternate names: Clara cell 10K secretory protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 20-Aug-1999
C:Accession: A36581; S10185; S21676
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta
J. Biol. Chem. 265, 12690-12693, 1990
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina
A:Reference number: A36581; MUID:90324266; PMID:2115524
A:Accession: A36581
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-96 <NOR>
A:Cross-references: GB:J05536; NID:9206039; PIDN:AAA41817.1; PID:9206040
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.
Nucleic Acids Res. 18, 2939-2946, 1990
A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region
A:Reference number: S10185; MUID:90272398; PMID:2349092
A:Accession: S10185
A:Status: transline expression not shown
A:Molecule type: DNA

Db 27 FORVIETLLMDTPSS-YEAMAELESPDQDMREAGAKLKLVDLTLPQKPREIIIMKEIA 85

Qy 67 GS 68
I

Db 86 QS 87

RESULT 4

S68231
FHG22 protein precursor - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 21-Aug-1998
C;Accession: S68231
R:Dominguez, P., 257-261, 1995
A;Title: Cloning of a Syrian hamster CDNA related to sexual dimorphism: establishment
A;Reference number: S68231; MUID:96105393; PMID:7498554
A;Accession: S68231
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-95 <DOM>
A:Cross-references: EMBL:266540
C:Superfamily: uteroglobin
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-95/Product: FHG22 protein #status predicted <MAT>

Query Match 17.9%; Score 76.5; DB 2; Length 95;
Best Local Similarity 37.9%; Pred. No. 0.27;
Matches 18; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

Qy 21 ELVEGTGLKYNVADKAAMTELKSCIDG-LQPMHKAELVKLLVOVLG 67
| | : | : | : | : | : | : | : | : | : | :
Db 41 KEFEAVLOTENAPPEAVKAVEKKCIDSTLNLYLEKMEMGKILAEVVG 88

RESULT 5

UGRB
uteroglobin precursor [validated] - rabbit
N:Alternate names: blastokinin
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Feb-1980 #sequence_revision 15-Oct-1982 #text_change 15-Sep-2000
C;Accession: A92391; A93461; A90935; A24217; A90303; A90417; A93824; A94608; I46904;
R:Bailly, A.; Agger, M.; Atger, P.; Carbon, M.A.; Allizon, M.; Vu Hai, M.T.; Logeat, F.
J. Biol. Chem. 258, 10384-10389, 1983
A;Title: The rabbit uteroglobin gene. Structure and interaction with the progesterone
A;Reference number: A92391; MUID:83290960; PMID:6309802
A;Accession: A92391
A:Molecule type: DNA
A:Residues: 1-91 <BAI>
A:Cross-references: GB:K00049; NID:g165789
R:Suske, G.; Wenz, M.; Cato, A.C.B.; Beato, M.
Nucleic Acids Res. 11, 2257-2271, 1983
A;Title: The uteroglobin gene region: hormonal regulation, repetitive elements and co
A;Reference number: A93461; MUID:83220783; PMID:630464
A;Accession: A93461
A:Molecule type: DNA
A:Residues: 1-91 <SUS>
A:Cross-references: GB:J00687; NID:g1772; PIDN:CAA25669.1; PID:g313668
R;Chandra, T.; Bullock, D.W.; Woo, S.L.C.
DNA 1, 19-26, 1981
A;Title: Hormonally regulated mammalian gene expression: steady-state level and nucle
A;Reference number: A90935; MUID:83157105; PMID:6299663
A;Accession: A90935
A:Molecule type: mRNA
A:Residues: 1-91 <CHA>
R:Lopez de Haro, M.S.; Nieto, A.
FEBS Lett. 193, 247-249, 1985
A;Title: Primary structure of rabbit lung uteroglobin as deduced from the nucleotide
A;Reference number: A24217; MUID:86056319; PMID:2415398
A;Accession: A24217
A:Molecule type: mRNA
A:Residues: 22-91 <LOP>

A;Cross-references: GB:M27564; NID:g165792; PIDN:AAA31496.1; PID:g165793
A;Experimental source: lung
R;Atger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.
Biochem. J. 177, 985-988, 1979
A;Title: N-terminal sequences of uteroglobin and its precursor.
A;Reference number: A90303; MUID:79187160; PMID:571719
A;Accession: A90303
A;Molecule type: protein
A;Residues: 1-5, 'F', 7-10, 'X', 15, 'G', 17-54, 'X', 56, 'B', 58-66, 'B', 68-70, 'XX', 73 <ATG>
R;Ponstingl, H.; Nicot, A.; Beato, M.
Biochemistry 17, 3908-3912, 1978
A;Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.
A;Reference number: A90417; MUID:79042086; PMID:568483
A;Accession: A90417
A;Molecule type: protein
A;Residues: 22-81, 'Q', 83-91 <PON>
R;Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978
A;Title: Amino acid sequence of a progesterone-binding protein.
A;Reference number: A93824; MUID:79074850; PMID:281700
A;Accession: A93824
A;Molecule type: protein
A;Residues: 22-49, 'D', 51, 'EN', 54-59, 61-66, 'NEPSL', 72-91 <POP>
R;Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
submitted to the Atlas, October 1982
A;Reference number: A94608
A;Accession: A94608
A;Molecule type: protein
A;Residues: 50-62; 67-71 <PO2>
R;Morize, I.; Surcouf, E.; Vaney, M.C.; Buehner, M.; Mornon, J.P.
submitted to the Brookhaven Protein Data Bank, April 1989
A;Reference number: A50025; PDB:1UTG
A;Contents: annotation: X-ray crystallography, 1.34 angstroms, residues 22-91
R;Morize, I.; Surcouf, E.; Vaney, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgrom, J.
J. Mol. Biol. 194, 725-739, 1987
A;Title: Refinement of the C22-1 crystal form of oxidized uteroglobin at 1.34 angstroms
A;Reference number: A44652; MUID:88011213; PMID:3656405
A;Contents: annotation: X-ray crystallography, 1.34 angstroms
R;Bally, R.; Delettre, J.
submitted to the Brookhaven Protein Data Bank, May 1989
A;Reference number: A50553; PDB:2UTG
A;Contents: annotation: X-ray crystallography, 1.64 angstroms, residues 22-91
R;Bally, R.; Delettre, J.
J. Mol. Biol. 206, 153-170, 1989
A;Title: Structure and refinement of the oxidized p2-1 form of uteroglobin at 1.64 angstroms
A;Reference number: A44653; MUID:89199637; PMID:2704039
A;Contents: annotation: X-ray crystallography, 1.64 angstroms; disulfide bonds
R;Menne, C.; Suske, G.; Arnemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982
A;Title: Isolation and structure of the gene for the progesterone-inducible protein uteroglobin
A;Reference number: 146904; MUID:83014990; PMID:6956897
A;Accession: 146904
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-91 <MEN>
A;Cross-references: GB:J00689; NID:g165786; PIDN:AAA31495.1; PID:g165788
R;Chandra, T.; Woo, S.L.C.; Bullock, D.W.
Biochem. Biophys. Res. Commun. 95, 197-204, 1980
A;Title: Cloning of the rabbit uteroglobin structural gene.
A;Reference number: 146905; MUID:81021016; PMID:7417250
A;Accession: 146905
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 39-77 <CH2>
A;Cross-references: GB:M25057; NID:g165802; PIDN:AAA31498.1; PID:g165803
R;Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.
Prog. Clin. Biol. Res. 85, 139-146, 1982
A;Title: Structure and regulated expression of the uteroglobin gene.
A;Reference number: 146907; MUID:82275176; PMID:6287481
A;Accession: 146907
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-45, 'V', 47-91 <SU2>

A;Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809
R;Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980
A;Title: Bacterial cloning of the rabbit uteroglobin structural gene.
A;Reference number: 146906; MUID:80241888; PMID:6156676
A;Accession: 146906
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 53-66, 'NT', 69-72 <AT2>
A;Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805
C;Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It is secreted by the uterus upon induction by progesterone.
C;Genetics:
A;Introns: 19/1; 81/3
C;Complex: homodimer linked by two disulfide bonds
C;Superfamily: uteroglobin
C;Keywords: homodimer; steroid binding; uterus
F;1-21/Domain: signal sequence #status experimental <SIG>
F;22-91/Product: uteroglobin #status experimental <MAT>
F;24/Disulfide bonds: Interchain (to 90) #status experimental
F;90/Disulfide bonds: Interchain (to 24) #status experimental
Query Match 17.4%; Score 74.5; DB 1; Length 91;
Best Local Similarity 27.4%; Pred. No. 0.43; Mismatches 17; Indels 1; Gaps 1;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
QY 7 FMDFLQTLVGTPEELVEGLKYNVNEADAKAMTELKSCIGLQPMHKAELVLLVQL 66
DB 27 FAHVIEENLLGTPTSS-YETSLKEFEDDTMKDAGQMKNVLDLSLPQTTREINIMLTKIV 85
QY 67 GS 68
DB 86 KS 87
RESULT 6
UGMS uteroglobin precursor - mouse
N;Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory protein
C;Species: Mus musculus (house mouse)
C;Date: 03-May-1994 #sequence, revision 21-Jan-1997 #text_change 22-Jun-1999
C;Accession: A53025; A56656; I51925; S24783
R;Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A;Title: Structure and regulation of the murine Clara cell secretory protein gene.
A;Reference number: A53025; MUID:94292183; PMID:8020953
A;Accession: A53025
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-96 <STR>
R;Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A;Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecular cloning
A;Reference number: A56656; MUID:93178380; PMID:8440203
A;Accession: A56656
A;Molecule type: mRNA; protein
A;Residues: 1-96 <SIN>
A;Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691
A;Experimental source: lung
A;Note: sequence extracted from NCBI backbone (NCBI:126148)
A;Note: parts of this sequence, including the amino end of the mature protein, were cDNA clones of the Clara cell secretory protein gene.
R;Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Nawkins, H.K.; DeMayo, F.J.
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A;Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell secretory protein
A;Reference number: I51925; MUID:94000840; PMID:8398159
A;Accession: I51925
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-96 <RES>
A;Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093
C;Genetics:
A;Introns: 19/1; 81/3
C;Complex: homodimer linked by two disulfide bonds

[illegible]

[illegible]

A:Residues: 1-91 <LOP>
A:Cross-references: GB:M25609; NID:g164246; PIDN:AAA30960.1; PID:g164247
A:Experimental source: lung
C:Comment: Uteroglobulin, synthesized in the uterus and lung, is secreted by the uterus up
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobulin
C:Keywords: lung; steroid binding; uterus
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-91/Product: uteroglobulin #status predicted <MAT>
F:24/Disulfide bonds: interchain (to 90) #status predicted
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 15.6% Score 66.5; DB 1; Length 91;
Best Local Similarity 25.8%; Pred. No. 3.3;
Matches 16; Conservative 17; Mismatches 28; Indels 1; Gaps 1;

QY 7 FMDLQTLVLGTPPEELYEGLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVKLLVQVL 66
DB 27 FAHVIENTLLGTPSS-YGTSLEKFPDDAMKDGQMKRVLDLTLPQTTRENIIRLTKIV 85
QY 67 GS 68
DB 86 KS 87

RESULT 15
AD3580
iron(III)-binding periplasmic protein precursor BMEII0565 [imported] - Brucella melitensis
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AD3580
R:DelVecchio, V.G.; Kapral, R.J.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.: Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AD3580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL53807.1; PID:g17984739; *GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEII0565
A:Map position: II

Query Match 15.6% Score 66.5; DB 2; Length 358;
Best Local Similarity 31.4%; Pred. No. 16;
Matches 22; Conservative 9; Mismatches 22; Indels 17; Gaps 3;

QY 22 LYEGTLG-KYVNEDAKAAMTELKSCIDGLQPMHKAEL-----VKLLVQVL 66
DB 142 VYAGALGWGYNTDIFAKNKLKPCWADLLDPYSYKGEIQMANPNSSGTYATLATLVQIM 201
QY 67 GSQDGAGTDY 76
DB 202 G--EDKGFDY 209

Search completed: April 28, 2003, 14:44:36
Job time : 18.7526 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:38:11 ; Search time 22.5155 Seconds
(without alignments)
444.049 Million cell updates/sec

Title: US-09-768-826-47_COPY_2_105
Perfect score: 537
Sequence: 1 KGRALLVALTLCICRMA.....VQVLGSDGAGTDKDDDK 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73.*
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	344.5	64.2	94	2 S17449	probable ligand-bi
2	96.5	18.0	96	2 A36581	polychlorinated bi
3	92	17.1	91	2 JS0036	Clara cell 10K pro
4	86	16.0	95	2 S68231	FHG22 protein prec
5	85	15.8	91	1 UGRB	uteroglobin precu
6	81.5	15.2	96	1 UGMS	cell specific 10K
7	78.5	14.6	113	2 JC2026	signal recognition
8	76.5	14.2	504	2 S05197	SRP 34 protein - h
9	76.5	14.2	504	2 S05198	uteroglobin precu
10	76.5	14.2	504	2 S54143	mannose-6-phosphat
11	73	14.2	91	1 UGRBL	probable polyketid
12	73	13.6	609	2 T38128	fibronectin-bindin
13	70.5	13.1	423	2 S41122	t/123.15 protein -
14	69.5	12.9	2458	2 T17420	hypothetical prote
15	69	12.8	540	2 G86790	hypothetical prote
16	68.5	12.8	2100	2 T38128	BN11 protein - yea
17	68	12.7	270	2 S30947	hypothetical prote
18	67.5	12.6	299	2 T16137	outer surface prot
19	67.5	12.6	1953	2 S63244	iron(III)-binding
20	67	12.5	511	2 T02269	sarcocystatin A pr
21	66.5	12.4	322	2 C83075	probable maltose/m
22	66.5	12.4	358	2 AD3580	3c3 20c protein -
23	66	12.3	192	2 S70285	phosphodiesterase
24	65.5	12.2	122	2 A43644	exoribonuclease, v
25	65.5	12.2	362	2 A71243	exoribonuclease, v
26	65.5	12.2	1321	2 T10929	exoribonuclease, v
27	65.5	12.2	1687	2 T30244	exoribonuclease, v
28	65.5	12.2	1706	2 T30175	exoribonuclease, v
29	65.5	12.2	1719	2 T30174	exoribonuclease, v

ALIGNMENTS

RESULT 1

S17449
Probable ligand-binding protein RYD5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S17449
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.
EMBO J. 10, 2813-2819, 1991

A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac
A:Reference number: S17447; MUID:92007724; PMID:1915264
A:Accession: S17449
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-94 <DE>
A:Cross-references: EMBL:X60661; NID:g57735; PIDN:CAA43068.1; PID:g57736

Query Match 64.2%; Score 344.5; DB 2; Length 94;
Best Local Similarity 75.0%; Pred. No. 2.1e-29;
Matches 69; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

Qy 1 KGRALLVALTLCICRMATGDNDEFFMDPLOTLLVGTPEELGKYNVNEDAKA 60
Db 2 KGSSA-LLVALTVLCIGLTRAEDNDEFFMDPLOTLLVGTPEELGKYNVNEDAKA 60
Qy 61 AMTELKSCIDGLQPMHKAELVLLVQVLGSDQ 92
Db 61 ALTELKSCIDELQPVHKEQLVLLVQVLDAQE 92

RESULT 2

A36581
Polychlorinated biphenyl-binding protein precursor - rat
N:Alternate names: Clara cell 10K secretory protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 20-Aug-1999
C:Accession: A36581; S10185; S21676
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta
J. Biol. Chem. 265, 12690-12693, 1990
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina
A:Reference number: A36581; MUID:90324266; PMID:2115524
A:Accession: A36581
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-96 <NOR>
A:Cross-references: GB:J05536; NID:g206039; PIDN:AAA41817.1; PID:g206040
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.
Nucleic Acids Res. 18, 2939-2946, 1990
A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region
A:Reference number: S10185; MUID:90272398; PMID:2349092
A:Accession: S10185
A:Status: translation not shown
A:Molecule type: DNA

Qy **5** ALLEVALTLFCICRMATGDEONDEFMDFLDTLLVGTPEELYEGTGLKYNVNEDAKAAMTE 64
 | : | : || | : | | ::|||: || | : : : : : | :

A:Reference number: A24217; MUID:86056319; PMID:2415398
A:Accession: A24217
A:Molecule type: mRNA
A:Residues: 22-91 <LOP>
A:Cross-references: GB:M27564; NID:gi65792; PIDN:AAA31496.1; PID:gi65793
A:Experimental source: lung
R:Atger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.
Biochem. J. 177, 985-988, 1979
A:Title: N-terminal sequences of uteroglobin and its precursor.
A:Reference number: A90303; MUID:79187160; PMID:571719
A:Accession: A90303
A:Molecule type: protein
A:Residues: 1-5, F', 7-10, X', 15, G', 17-54, X', 56, B', 58-66, B', 68-70, XX', 73 <ATG>
R:Ponstingl, H.; Nieto, A.; Beato, M.
Biochemistry 17, 3908-3912, 1978
A:Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.
A:Reference number: A90417; MUID:79042086; PMID:568483
A:Accession: A90417
A:Molecule type: protein
A:Residues: 22-81, Q', 83-91 <PON>
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978
A:Title: Amino acid sequence of a progesterone-binding protein.
A:Reference number: A93824; MUID:79074850; PMID:281700
A:Accession: A93824
A:Molecule type: protein
A:Residues: 22-49, D', 51, EN', 54-59, 61-66, NEPSL', 72-91 <POP>
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
submitted to the Atlas, October 1982
A:Reference number: A94608
A:Accession: A94608
A:Molecule type: protein
A:Residues: 50-62, 67-71 <PO2>
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Buehner, M.; Mornon, J.P.
submitted to the Brookhaven Protein Data Bank, April 1989
A:Reference number: A50025; PDB:1UTG
A:Contents: annotation; X-ray crystallography, 1.34 angstroms, residues 22-91
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgrom, J.
J. Mol. Biol. 194, 725-739, 1987
A:Title: Refinement of the C22-1 crystal form of oxidized uteroglobin at 1.34 angstroms
A:Reference number: A4652; MUID:88011213; PMID:3656405
A:Contents: annotation; X-ray crystallography, 1.34 angstroms
R:Bally, R.; Delettre, J.
submitted to the Brookhaven Protein Data Bank, May 1989
A:Reference number: A50553; PDB:2UTG
A:Contents: annotation; X-ray crystallography, 1.64 angstroms, residues 22-91
R:Bally, R.; Delettre, J.
J. Mol. Biol. 206, 153-170, 1989
A:Title: Structure and refinement of the oxidized P2-1 form of uteroglobin at 1.64 angstroms
A:Reference number: A44653; MUID:89199637; PMID:2704039
A:Contents: annotation; X-ray crystallography, 1.64 angstroms; disulfide bonds
R:Menne, C.; Suske, G.; Arneemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982
A:Title: Isolation and structure of the gene for the progesterone-inducible protein uteroglobin
A:Reference number: I46904; MUID:83014990; PMID:6956897
A:Accession: I46904
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-91 <MEN>
R:Chandra, T.; Woo, S.L.C.; Bullock, D.W.
Biochem. Biophys. Res. Commun. 95, 197-204, 1980
A:Title: Cloning of the rabbit uteroglobin structural gene.
A:Reference number: I46905; MUID:81021016; PMID:7417250
A:Accession: I46905
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 39-77 <CH2>
A:Cross-references: GB:M25057; NID:gi165802; PIDN:AAA31498.1; PID:gi165803
R:Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.
Prog. Clin. Biol. Res. 85, 139-146, 1982
A:Title: Structure and regulated expression of the uteroglobin gene.
A:Reference number: I46907; MUID:82275176; PMID:6287481

A:Accession: I46907
A:A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-45,'V',47-91 <SU2>
A:Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809
R:Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980
A>Title: Bacterial cloning of the rabbit uteroglobin structural gene.
A:Reference number: I46906; MUID:80241888; PMID:6156676
A:Accession: I46906
A:A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 53-66,'NT',69-72 <AT2>
A:Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805
C:Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It b
C:Genetics:
A:Introns: 19/1: 81/3
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobulin
C:Keywords: homodimer; steroid binding; uterus
F:1-21/DNA: signal sequence #status experimental <SIG>
F:22-91/Product: uteroglobulin #status experimental <MAT>
F:24/Disulfide bonds: interchain (to 90) #status experimental
F:90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 15.8%; Score 85; DB 1; Length 91;
Best Local Similarity 27.5%; Pred. No. 0.066;
Matches 25; Conservative 17; Mismatches 33; Indels 16; Gaps 3;

Qy 6 LLLVALFLFC-----ICRMATGEDNEFFMDPLOTLLVGTPPELYEGTIGKYNVEDAK 59
||| || | ||
Db 7 LAVTLALLCSPASAGCP-----FAHVNIENLLGTGPS-YETSLKEFPDPTMK 56

Qy 60 AAMTELKSCIDGLQPMHKAELVKLVQLGS 90
| : : | : | : : : : :
Db 57 DAGQMQRKVLDSPQTREINIMKLTERIKVS 87

RESULT 6
UGMS
uteroglobin precursor - mouse
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory prote
C:Species: Mus musculus (house mouse)
C>Date: 03-May-1994 #sequence_revision 21-Jan-1997 #text_change 22-Jun-1999
C:Accession: A53025; A56656; I51925; S24783
R:Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A>Title: Structure and regulation of the murine Clara cell secretory protein gene.
A:Reference number: A53025; MUID:94292183; PMID:8020953
A:Accession: A53025
A:A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <STR>
A:Cross-references: GB:I24372; NID:g461147; PIDN:AA65446.1; PID:g785054
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A>Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecu
A:Reference number: A56656; MUID:93178380; PMID:8440203
A:Accession: A56656
A:Molecule type: mRNA
A:Residues: 1-96 <SIN>
A:Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691
A:Experimental source: lung
A>Note: sequence extracted from NCBI backbone (NCBIP:126148)
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Hawkins, H.K.; DeMayo, F.J.
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A>Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell
A:Reference number: I51925; MUID:94000840; PMID:8398159
A:Accession: I51925
A:A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-96 <RES>

A:Accession: S10619
A:Molecule type: protein
A:Residues: 2-35;58-70;120-126;141-153;374-388;416-426;470-485 <BER>
C:Superfamily: signal recognition particle 54K protein
C:Keywords: GTP binding; nucleotide binding; P-loop
F:1-504/Product: signal recognition particle 54K protein #status experimental <MAT>
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match
Best Local Similarity 14.2%; Score 76.5; DB 2; Length 504;
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 1 KGSRALLLVAULTFCICRMATGEDNDEF----FMDFLQTL--VGTPEELYEGTLGKYNVN 55
DB 255 KGGGSAVAATKSPFIIFGTGHEIHDDPEPKTQPFISKLLGMGDIEGLID-----KVN 308

QY 56 E----DAKAAMTELKSCIDIGLOPMHK--AELVKL--LVQVLGSDGAGTDDY 98
DB 309 ELKDDNEALIEKLKHGQFTLRDMYEQFQINMKMGPFQILGMIPFGGTFD 359

RESULT 9
S05198
signal recognition particle 54K protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999
C:Accession: S05198; S14553
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.
Nature 340, 482-486, 1989
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit
A:Reference number: S05198; MUID:89330597; PMID:2502718
A:Accession: S05198
A:Molecule type: mRNA
A:Residues: 1-504 <BER1>
A:Cross-references: EMBL:X16319
R:Bernstein, H.
submitted to the EMBL Data Library, November 1989
A:Reference number: S14553
A:Accession: S14553
A:Molecule type: mRNA
A:Residues: 1-42, 'DV', 45-504 <BER2>
A:Cross-references: EMBL:X16319; NID:954193; PIDN:CAA34386.1; PID:g54194
C:Superfamily: signal recognition particle 54K protein
F:2-504/Product: signal recognition particle 54K protein #status predicted <MAT>

Query Match
Best Local Similarity 14.2%; Score 76.5; DB 2; Length 504;
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 1 KGSRALLLVAULTFCICRMATGEDNDEF----FMDFLQTL--VGTPEELYEGTLGKYNVN 55
DB 255 KGGGSAVAATKSPFIIFGTGHEIHDDPEPKTQPFISKLLGMGDIEGLID-----KVN 308

QY 56 E----DAKAAMTELKSCIDIGLOPMHK--AELVKL--LVQVLGSDGAGTDDY 98
DB 309 ELKDDNEALIEKLKHGQFTLRDMYEQFQINMKMGPFQILGMIPFGGTFD 359

RESULT 10
S54143
SRP 54 protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 23-Jul-1999
C:Accession: S54143
R:Patel, S.
submitted to the EMBL Data Library, April 1995
A:Reference number: S54143
A:Accession: S54143
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-504 <PAT>
A:Cross-references: EMBL:X86373; NID:g784998; PIDN:CAA60132.1; PID:g784999
C:Superfamily: signal recognition particle 54K protein

```

Query Match          14.2%; Score 76.5; DB 2; Length 504;
Best Local Similarity 32.4%; Pred. No. 3.6;
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

      2Y      1 KGSRALLVALTFCICRMATGEDNDEF---FMDFLOTLL-VGTPPEELVEGTLGKYNVN 55
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
      Db      255 KGGGALSAAVAATKSPFIIFGTGEGHIDDFEPFKTOPFISKLLGMGDIEGLID-----KVN 308

      2Y      56 E-----DAKAAMTELSCIDGLQPMHK--AELVKL--LVQVLGSDQGAGTDY 98
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
      Db      309 ELKLDNDEALTEKLGKHQGFOTLRDMYEQFNIMKMGPFQSILGMPFGTDF 359

RESULT 11
JGRBL
uteroglobin precursor - brown hare
N:Alternate names: blastokinin
S:Species: Lepus capensis (brown hare)
D:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999

```

RESULT 11
 UGBL
 uterogloblin precursor - brown hare
 Alternate names: blastokinin
 Species: Lepus capensis (brown hare)
 Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999
 Accession: A23825
 Lopez de Haro, M.S.; Nieto, A.
 Biochem. J. 233, 895-898, 1986
 Title: Nucleotide and derived amino acid sequences of a cDNA coding for pre-uterogloblin
 Reference number: A23825; MUID:86323069; PMID:3019311
 Accession: A23825
 Molecule type: mRNA
 Residues: 1-91 <LOP>
 Cross-references: GB:M25609; NID:gl64246; PIDN:AAA30960.1; PID:gl64247
 Experimental source: lung
 Comment: Uterogloblin, synthesized in the uterus and lung, is secreted by the uterus up
 Complex: homodimer linked by two disulfide bonds
 Superfamily: uterogloblin
 Keywords: lung; steroid binding; uterus
 1-21/Domain: signal sequence #status predicted <STC>
 22-91/Product: uterogloblin #status predicted <WAT>
 24/disulfide bonds: interchain (to 90) #status predicted
 90/disulfide bonds: interchain (to 24) #status predicted

```

Query Match          14.2%  Score 76;  DB 1;  Length 91;
Best Local Similarity 26.4%;  Pred.No. 0.59;
Matches 24;  Conservative 17;  Mismatches 34;  Indels 16;  Gaps 3;

6  LLLVALTLCFC-----ICRMATGDENDFFDFLQTLVLVGTPEELYEGTIGKYNVEDAK 59
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
7  LALVTALLCSPASAGIC-----PGFAHVIENLLLTGPSS-YGTSLKEFPQDDAMK 56

60 AAMTELKSCIDGLQPMHRAELVLLVQVLGS 90
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
57 DAGMOMKVLDTLPOTTRENIILKTEKIVKS 87
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

RESULT 12
 T38656
 Probable RNA-binding protein - fission yeast (Schizosaccharomyces pombe)
 ;Species: Schizosaccharomyces pombe
 ;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 ;Accession: T38656
 ;Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 Submitted to the EMBL Data Library, August 1997
 ;Reference number: Z21804
 ;Accession: T38656
 ;Status: preliminary; translated from GB/EMBL/DDBJ
 ;Molecule type: DNA
 ;Residues: 1-609 <MUR>
 ;Cross-references: EMBL:AL109739; NID:el534774; PIDN:CAB52270.1; GSPDB:GN000066; SPDB:SH
 ;Experimental source: strain 972h.; cosmid c343
 ;Genetics:
 ;Gene: SPDB:SPAC343.07
 ;Map position: 1
 ;Introns: 44/1; 96/2; 325/3

Query Match 13.6%; Score 73; DB 2; Length 609;

```

Best_Local_Similarity 23.58; Pred. No. 10;
Matches 23; Conservative 20; Mismatches 27; Indels 28; Gaps 5;

Qy 17 CRMATG---ENDENFFMDFLOTLVG-----TPELYE-----GTLGKYNVNE 56
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 248 CQMPGMLLEQNPQFLVDNPSIFVIGILNLPLKVKSPVELYNFESNHGHILGVAINQSINE 307
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 57 D---AKAAATLEKSIDGLQPMH-----KAEVLKLLVQ 86
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 308 DMTHYAEVAVVSTYESCIETIEKFAHIAVEGSIQLQLFK 345
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 13
S41122
mannose-6-phosphate isomerase (EC 5.3.1.8) - human
N:Alternate names: phosphomannose isomerase
C:Species: Homo sapiens (man)
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C:Accession: S41122; S38666
R:Proudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.
Eur. J. Biochem. 219, 415-423, 1994
A:Title: Purification, cDNA cloning and heterologous expression of human phosphomannose
A:Reference number: S41122; MUID:94139717; PMID:8307007

```

RESULT 13
S41122
mannose-6-phosphate isomerase (EC 5.3.1.8) - human
N:Alternate names: phosphomannose isomerase
C:Species: Homo sapiens (man)
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 08-Oct-1999
C:Accession: S41122; S38666
R:Proudfoot, A.E.I.; Turcatti, G.; Wells, T.N.C.; Payton, M.A.; Smith, D.J.
Eur. J. Biochem. 219, 415-423, 1994
A:Title: Purification, cDNA cloning and heterologous expression of human phosphomanno
A:Reference number: S41122; MUID:94139717; PMID:8307007
A:Accession: S41122
A:Molecule type: mRNA
A:Residues: 1-423 <PRO>
A:Cross-references: EMBL:X76057; NID:g416016; PIDN:CAA53657.1; PID:g416017
C:Genetics:
A:Gene: GDB:MFI
A:Cross-references: GDB:119397; OMIM:154550
A:Map position: 15q22-15qter
C:Superfamily: yeast mannose-6-phosphate isomerase
C:Keywords: intramolecular oxidoreductase; isomerase

```
Query Match      13.1%; Score 70.5; DB 2; Length 423;  
Best Local Similarity 25.7%; Pred. No.13;  
Matches 27; Conservative 19; Mismatches 38; Indels 21; Gaps 5;
```

Oy 7 LLVALTUF-CICRMATGEDNDEFFMDFLOTLLVGTPPEELYEGTLG-----KYNVNEDA 58
Db : ||| | | | : | : | : | : | :
Oy 138 MAIATLFPQGLCGRPVEE-----IVFLKKVP--FQLIGDEAATHLKQTMSHDS 187
Db : ||| | | | : | : | : | : | :

Oy 59 KAAMTELKSCIDGIQPMMHK --- AELVKLLVVQLGSQDGAGTDYKD 100
Db : |||| : ||| : ||| : ||| :

Oy 188 QAVASSIQSGCFSLMKSEKVWVLEOLNLVLKRISOOAAACGNMED 232
Db : |||| : ||| : ||| : ||| :

RESULT 14
T17420
probable polyketide synthase type I - *Pseudomonas fluorescens*
C:Species: *Pseudomonas fluorescens*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C:Accession: T17420
E:Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Loper, J.E.
J. Bacteriol. 181, 2166-2174, 1999
A:Title: Characterization of the pyoluteorin biosynthetic gene cluster of *Pseudomonas*
A:Reference number: Z18776; MUID:99194726; PMID:10094695
A:Accession: T17420
A:Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-2458 <NOW>
A:Cross-references: EMBL:AF081920; NID:g4582974; PID:g2781416; PIDN:AAC38075.1
C:Genetics:
A:Gene: pltB
C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier pro
C:Keywords: carrier protein
F:31-429/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
F:535-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>
F:939-1009/Domain: acyl carrier protein homology <ACP1>
F:1053-1446/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
F:2337-2408/Domain: acyl carrier protein homology <ACP2>

Query Match	Score 69.5;	DB 2;	Length 2458;
12.98;			

```
Best Local Similarity 32.1%; Pred. No. 1.2e+02;
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;

QY 43 ELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHRAELVKLVQVLGSDGAG 95
      :| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2239 QLLHLTQSELEQDA-AARTAMQG--SGLQLPQRSQIVQAIARVLGGGQCG 2288

RESULT 15
G86790
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86790
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86790
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-540 <STO>
A:Cross-references: GB:AF005176; PID:g12724308; PIDN:AAK05425.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yngB

Query Match 12.8%; Score 69; DB 2; Length 540;
Best Local Similarity 30.6%; Pred. No. 24;
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 24 DNDEFFM---DFLQTLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHRAE 79
      || | | | | | | | | | | | | | | | | | | | | | | | |
DB 294 DNAEIFRQKGLLNTFLNQVPNDKTSVTLENYVTNEPIELNLPALSPVQNAQRVFRHYQ 353

QY 80 LVKLLVQVLGSQ 91
      :| | | | |
DB 354 KLAQAVKFLGEQ 365

Search completed: April 28, 2003, 14:44:37
Job time : 23.5155 secs
```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: April 28, 2003, 14:38:11 ; Search time 22.732 seconds
(without alignments)
444.049 Million cell updates/sec

Title: US-09-768-826-47
Perfect score: 542
Sequence: 1 MKGSRALLLVLTLCICRM.....VQVLGSQDAGCTKYKDDDDK 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	349.5	64.5	94	2	probable ligand-bi
2	96.5	17.8	96	2	polychlorinated bi
3	92	17.0	91	2	Clara cell 10K pro
4	91	16.8	95	2	FHG22 protein prec
5	89	16.4	91	1	uteroglobin precu
6	81.5	15.0	96	1	uteroglobin precu
7	80	14.8	91	1	uteroglobin precu
8	78.5	14.5	113	2	cell specific 10K
9	76.5	14.1	504	2	signal recognition
10	76.5	14.1	504	2	signal recognition
11	76.5	14.1	504	2	SRP 54 protein - h
12	73	13.5	609	2	probable RNA-bind
13	70.5	13.0	423	2	mannose-6-phosphat
14	69.5	12.8	2458	2	probable polyketid
15	69	12.7	540	2	fibronectin-bindin
16	68.5	12.6	2100	2	t123.15 protein -
17	68	12.5	92	2	major allergen Fel
18	68	12.5	270	2	hypothetical prote
19	67.5	12.5	299	2	hypothetical prote
20	67.5	12.5	322	2	octaprenyl-diphosp
21	67.5	12.5	1553	2	BNII protein - yea
22	67	12.4	511	2	hypothetical prote
23	66.5	12.3	358	2	iron(III)-binding
24	66	12.2	192	2	outer surface prot
25	65.5	12.1	92	2	major allergen cha
26	65.5	12.1	122	2	sarcocystatin A pr
27	65.5	12.1	362	2	probable maltose/m
28	65.5	12.1	1321	2	3C3.20c protein -
29	65.5	12.1	1687	2	phosphodiesterase

30	65.5	12.1	1706	2	T30175
31	65.5	12.1	1719	2	T30174
32	65	12.0	209	2	S73690
33	65	12.0	255	2	B64175
34	65	12.0	312	2	G71146
35	65	12.0	842	2	E96641
36	65	12.0	2233	1	ZLNZP3
37	64.5	11.9	339	2	E90436
38	64.5	11.9	403	2	B89808
39	64.5	11.9	454	1	XNRTY
40	64.5	11.9	518	2	H85440
41	64.5	11.9	882	1	WMBEU5
42	64	11.8	190	2	T20725
43	64	11.8	342	1	C69395
44	64	11.8	446	2	G81326
45	64	11.8	627	2	H96951

ALIGNMENTS

RESULT 1

SI7449
probable ligand-binding protein RYD5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: SI7449
R:Dear, T.N.; Boehm, T.; Keverne, E.B.; Rabbitts, T.H.
EMBO J. 10, 2813-2819, 1991
A:Title: Novel genes for potential ligand-binding proteins in subregions of the olfac
A:Reference number: SI7447; MUID:92007724; PMID:1915264
A:Accession: SI7449
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-94 <DEA>
A:Cross-references: EMBL:X60661; NID:g57735; PIDN:CRA43068.1; PID:g57736

Query Match 64.5%; Score 349.5; DB 2; Length 94;
Best Local Similarity 75.3%; Pred. No. 7.3e-30;
Matches 70; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY	1	MKGSRALLLVLTLCICRMATGEDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAK 60
DB	1	MKGSSA-LLVALTLCICGLTAEEDNDEFFMDFLQTLVGTPEELYEGTLGKYNVEDAK 59
QY	61	AAATELKSCIDGLQPMHKAELVQLVQVLSQD 93
DB	60	AAATELKSCIDELQPVHKEQLVQLVQVLSQD 92

RESULT 2

A36581
polychlorinated biphenyl-binding protein precursor - rat
N:Alternate names: Clara cell 10K secretory protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 08-Mar-1991 #sequence_revision 08-Mar-1991 #text_change 20-Aug-1999
C:Accession: A36581; S10185; S21676
R:Nordlund-Moeller, L.; Andersson, O.; Ahlgren, R.; Schilling, J.; Gillner, M.; Gusta
J. Biol. Chem. 265, 12690-12693, 1990
A:Title: Cloning, structure, and expression of a rat binding protein for polychlorina
A:Reference number: A36581; MUID:90324266; PMID:2115524
A:Accession: A36581
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-96 <NOR>
A:Cross-references: GB:J05536; NID:Q206039; PIDN:AAA41817.1; PID:Q206040
R:Hagen, G.; Wolf, M.; Katyal, S.L.; Singh, G.; Beato, M.; Suske, G.
Nucleic Acids Res. 18, 2939-2946, 1990
A:Title: Tissue-specific expression, hormonal regulation and 5'-flanking gene region
A:Reference number: S10185; MUID:90272398; PMID:2349092
A:Accession: S10185
A:Status: translation not shown
A:Molecule type: DNA

10

A:Reference number: A24217; MUID:86056319; PMID:2415398
A:Accession: A24217
A:Molecule type: mRNA
A:Residues: 22-91 <LOP>
A:Cross-references: GB:M27564; NID:g165792; PIDN:AAA31496.1; PID:g165793
A:Experimental source: lung
R:Atger, M.; Mercier, J.C.; Haze, G.; Fridlansky, F.; Milgrom, E.
Biochem. J. 177, 985-988, 1979
A:Title: N-terminal sequences of uteroglobin and its precursor.
A:Reference number: A90303; MUID:79187160; PMID:571719
A:Accession: A90303
A:Molecule type: protein
A:Residues: 1-5, F',7-10, X',15, G',17-54, X',56, B',58-66, B',68-70, XX',73 <ATG>
R:Ponstingl, H.; Nieto, A.; Beato, M.
Biochemistry 17, 3908-3912, 1978
A:Title: Amino acid sequence of progesterone-induced rabbit uteroglobin.
A:Reference number: A90417; MUID:79042086; PMID:568483
A:Accession: A90417
A:Molecule type: protein
A:Residues: 22-81, Q',83-91 <PON>
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
Proc. Natl. Acad. Sci. U.S.A. 75, 5516-5519, 1978
A:Title: Amino acid sequence of a progesterone-binding protein.
A:Reference number: A93824; MUID:79074850; PMID:281700
A:Accession: A93824
A:Molecule type: protein
A:Residues: 22-49, D',51, EN',54-59,61-66, NEPSL',72-91 <POP>
R:Popp, R.A.; Foresman, K.R.; Wise, L.D.; Daniel Jr., J.C.
submitted to the Atlas, October 1982
A:Reference number: A94608
A:Accession: A94608
A:Molecule type: protein
A:Residues: 50-62;67-71 <PO2>
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Buehner, M.; Mornon, J.P.
submitted to the Brookhaven Protein Data Bank, April 1989
A:Reference number: A50025; PDB:1UTG
A:Contents: annotation; X-ray crystallography, 1.34 angstroms, residues 22-91
R:Morize, I.; Surcouf, E.; Vanev, M.C.; Epelboin, Y.; Buehner, M.; Fridlansky, F.; Milgrom, J.
J. Mol. Biol. 194, 725-739, 1987
A:Title: Refinement of the C222-1 crystal form of oxidized uteroglobin at 1.34 angstroms
A:Reference number: A4652; MUID:88011213; PMID:3656405
A:Contents: annotation; X-ray crystallography, 1.34 angstroms
R:Bally, R.; Delettre, J.
submitted to the Brookhaven Protein Data Bank, May 1989
A:Reference number: A50553; PDB:2UTG
A:Contents: annotation; X-ray crystallography, 1.64 angstroms, residues 22-91
R:Bally, R.; Delettre, J.
J. Mol. Biol. 206, 153-170, 1989
A:Title: Structure and refinement of the oxidized P2-1 form of uteroglobin at 1.64 angstroms
A:Reference number: A44653; MUID:89199637; PMID:2704039
A:Contents: annotation; X-ray crystallography, 1.64 angstroms; disulfide bonds
R:Menne, C.; Suske, G.; Arnemann, J.; Wenz, M.; Cato, A.C.B.; Beato, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 4853-4857, 1982
A:Title: Isolation and structure of the gene for the progesterone-inducible protein uteroglobin
A:Reference number: I46904; MUID:83014990; PMID:6956897
A:Accession: I46904
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-91 <MEN>
A:Cross-references: GB:J00689; NID:g165786; PIDN:AAA31495.1; PID:g165788
R:Chandra, T.; Woo, S.L.C.; Bullock, D.W.
Biochem. Biophys. Res. Commun. 95, 197-204, 1980
A:Title: Cloning of the rabbit uteroglobin structural gene.
A:Reference number: I46905; MUID:81021016; PMID:7417250
A:Accession: I46905
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 39-77 <CH2>
A:Cross-references: GB:M25057; NID:g165802; PIDN:AAA31498.1; PID:g165803
R:Suske, G.; Menne, C.; Cato, A.; Wenz, M.; Beato, M.
Prog. Clin. Biol. Res. 85, 139-146, 1982
A:Title: Structure and regulated expression of the uteroglobin gene.
A:Reference number: I46907; MUID:82275176; PMID:6287481

A:Accession: I46907
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-45, 'V', 47-91 <SU2>
A:Cross-references: GB:M32012; NID:g165807; PIDN:AAA31500.1; PID:g165809
R:Atger, M.; Perricaudet, M.; Tiollais, P.; Milgrom, E.
Biochem. Biophys. Res. Commun. 93, 1082-1088, 1980
A:Title: Bacterial cloning of the rabbit uteroglobin structural gene.
A:Reference number: I46906; MUID:80241888; PMID:6156676
A:Accession: I46906
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 53-66, 'NT', 69-72 <AT2>
A:Cross-references: GB:M25038; NID:g165804; PIDN:AAA31499.1; PID:g165805
C:Comment: Uteroglobin is secreted by the uterus upon induction by progesterone. It b
C:Genetics:
A:Introns: 19/1; 81/3
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobin
C:Keywords: homodimer; steroid binding; uterus
F:1-21/DNA: signal sequence #status experimental <SIG>
F:22-91/Product: uteroglobin #status experimental <MAT>
E:24/Disulfide bonds: interchain (to 90) #status experimental
F:90/Disulfide bonds: interchain (to 24) #status experimental

Query Match 16.4%; Score 89; DB 1; Length 91;
Best Local Similarity 27.8%; Pred. No. 0.026;
Matches 27; Conservative 18; Mismatches 36; Indels 16; Gaps 3;

Qy 1 MKGSRALLLVALTFC-----ICRMATGENDFFMDFLOTLLVTGPPELYEGTLGKYN 54
|| | ||| ||| || || : : ||||| || : :
Db 1 MKLAITLAVTLALCSPASAGICPR-----FAHVIENTLLGTPTSS-YETSLKEFE 50
:: | | :: | | : : : : : : : : : : :

Qy 55 VNEDAKAAWTELKSCIDGLOPMHKAELVKLLVQLGS 91
:: | | :: | | : : : : : : : : : : :

Db 51 PDPTMKDAGMQKKVLDSLPLPOTREINMKUTEIKVS 87
:: | | :: | | : : : : : : : : : : :

RESULT 6
UGMS
uteroglobin precursor - mouse
N:Alternate names: CC10; Clara cell 10K protein precursor; Clara cell secretory prote
C:Species: Mus musculus (house mouse)
C>Date: 03-May-1994 #sequence.revision 21-Jan-1997 #text_change 22-Jun-1999
C:Accession: A53025; A56656; I51925; S24783
R:Stripp, B.R.; Huffman, J.A.; Bohinski, R.J.
Genomics 20, 27-35, 1994
A:Title: Structure and regulation of the murine Clara cell secretory protein gene.
A:Reference number: A53025; MUID:94292183; PMID:8020953
A:Accession: A53025
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <ST>
A:Cross-references: GB:I24372; NID:g461147; PIDN:AAA65446.1; PID:g785054
R:Singh, G.; Katyal, S.L.; Brown, W.E.; Kennedy, A.L.
Exp. Lung Res. 19, 67-75, 1993
A:Title: Mouse Clara cell 10-kDa (CC10) protein: cDNA nucleotide sequence and molecu
A:Reference number: A56656; MUID:93178380; PMID:8440203
A:Accession: A56656
A:Molecule type: mRNA
A:Residues: 1-96 <SIN>
A:Cross-references: EMBL:X67702; NID:g49690; PIDN:CAA47936.1; PID:g49691
A:Experimental source: lung
A>Note: sequence extracted from NCBI backbone (NCBIP:126148)
R:Margraf, L.R.; Finegold, M.J.; Stanley, L.A.; Major, A.; Nawkins, H.K.; DeMayo, F.J.
Am. J. Respir. Cell Mol. Biol. 9, 231-238, 1993
A:Title: Cloning and tissue-specific expression of the cDNA for the mouse Clara cell
A:Reference number: I51925; MUID:94000840; PMID:8398159
A:Accession: I51925
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-96 <RES>

A:Cross-references: GB:L04503; NID:g202313; PIDN:AAA03625.1; PID:g433093
A:Genetics: 19/1; 81/3
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobin
C:Keywords: lung; steroid binding; uterus
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-96/Product: uteroglobin #status predicted <MAT>
F:24/Disulfide bonds: interchain (to 90) #status predicted
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 15.0%; Score 81.5; DB 1; Length 96;
Best Local Similarity 26.2%; Pred. No. 0.17;
Matches 22; Conservative 19; Mismatches 42; Indels 1; Gaps 1;

QY 8 LLVALTLCICRMATGEDNDEFMFLOTLVLTPEELVGTGKYNVNEAKAAMTELK 67
DB 5 ITITVMSLICSSASSSDICPFLQVLEALLMES-ESGYVASLKPFPNPGSDLNAGTQLK 63

QY 68 SCIDGLQPMHKAELVKLLVQVLGS 91
DB 64 RLVDTLPQETRINIMKLEIKILTS 87

RESULT 7
UGRBL
uteroglobin precursor - brown hare
N:Alternate names: blastokinin
C:Species: Lepus capensis (brown hare)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 22-Jun-1999
C:Accession: A23825
R:Lopez de Haro, M.S.; Nieto, A.
Biochem. J. 235, 895-898, 1986
A:Title: Nucleotide and derived amino acid sequences of a cDNA coding for pre-uteroglobin
A:Reference number: A23825; MUID:86323069; PMID:3019311
A:Accession: A23825
A:Molecule type: mRNA
A:Residues: 1-91 <LOP>
A:Cross-references: GB:M25609; NID:g164246; PIDN:AAA30960.1; PID:g164247
A:Experimental source: lung
C:Comment: Uteroglobin, synthesized in the uterus and lung, is secreted by the uterus up
C:Complex: homodimer linked by two disulfide bonds
C:Superfamily: uteroglobin
C:Keywords: lung; steroid binding; uterus
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-91/Product: uteroglobin #status predicted <MAT>
F:24/Disulfide bonds: interchain (to 90) #status predicted
F:90/Disulfide bonds: interchain (to 24) #status predicted

Query Match 14.8%; Score 80; DB 1; Length 91;
Best Local Similarity 26.8%; Pred. No. 0.23;
Matches 26; Conservative 18; Mismatches 37; Indels 16; Gaps 3;

QY 1 MKSRALLVLTLCFCICRMATGEDNDEFMFLOTLVLTPEELVGTGKYNV 54
DB 1 MKLTITLALTALICLSPASAGIC-----PGFAHVHLLGTPTSS-YGTSLKBEFQ 50

QY 55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGS 91
DB 51 PDDAMKAGMQMKVLDLTLPQTTRENIIMKLEIKIVKS 87

RESULT 8
JC2026
cell specific 10K protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 17-Mar-1999
C:Accession: JC2026
R:Ray, M.K.; Magdalen, S.; O'Malley, B.W.; DeMayo, F.J.
Biochem. Biophys. Res. Commun. 197, 163-171, 1993
A:Title: Cloning and characterization of the mouse clara cell specific 10 kDa protein ge
A:Reference number: JC2026; MUID:94071937; PMID:7916613
A:Accession: JC2026

A:Molecule type: DNA
A:Residues: 1-113 <RAY>
C:Comment: This protein is the major secretory product of the Clara cell and binds to
C:Superfamily: uteroglobin
F:73/Region: ochre stop codon

Query Match 14.5%; Score 78.5; DB 2; Length 113;
Best Local Similarity 30.7%; Pred. No. 0.43;
Matches 27; Conservative 15; Mismatches 43; Indels 3; Gaps 2;

QY 6 ALLVALTLCICRMATGEDND--EFFMDFLOTLVLTPEELVGTGKYNVNEAKAAM 63
DB 18 SLEKKLLSFCFLPATPASSDICPFLQVLEALLMES-ESGYVASLKPFPNPGSDLNAG 76

QY 64 TELKSCIDGLQPMHKAELVKLLVQVLGS 91
DB 77 TOLKRLVDTLPQETRINIMKLEIKILTS 104

RESULT 9
S05197
signal recognition particle 54K protein - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Feb-2001
C:Accession: S05197; S10619
R:Roemisch, K.; Webb, J.; Herz, J.; Prehn, S.; Frank, R.; Vingron, M.; Dobberstein, B.
Nature 340, 478-482, 1989
A:Title: Homology of 54K protein of signal-recognition particle, docking protein and
A:Reference number: S05197; MUID:89330596; PMID:2502717
A:Accession: S05197
A:Molecule type: mRNA
A:Residues: 1-504 <ROE>
A:Cross-references: EMBL:X16318; NID:g928; PIDN:CAA34385.1; PID:g930
A:Note: part of this sequence, including the amino end of the mature protein, was con
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.
Nature 340, 482-486, 1989
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit
A:Reference number: S05198; MUID:89330597; PMID:2502718
A:Accession: S10619
A:Molecule type: protein
A:Residues: 2-35;58-70;120-126;141-153;374-388;416-426;470-485 <BER>
C:Superfamily: signal recognition particle 54K protein
C:Keywords: GTP binding; nucleotide binding; P-loop
F:1-504/Product: signal recognition particle 54K protein #status experimental <MAT>
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 14.1%; Score 76.5; DB 2; Length 504;
Best Local Similarity 32.4%; Pred. No. 3.7;
Matches 36; Conservative 16; Mismatches 40; Indels 19; Gaps 6;

QY 2 KGSRALLVALTLCFCICRMATGEDNDEF----FMDFLQTLVLTPEELVGTGKYNV 56
DB 255 KGGGALSAAVATKSPFIIFIGTGEHIDREPPKTPQPFISKLLGMGIDGLID-----KVN 308

QY 57 E----DAKAAMTELKSCIDGLQPMHK--AELVKL--LVQVLGSQDCAGTDY 99
DB 309 ELKLDNEALIEKLRHGQFTLRDMYEQFNIMKMGPFQILGMIPGFGTDF 359

RESULT 10
S05198
signal recognition particle 54K protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 23-Jul-1999
C:Accession: S05198; S14553
R:Bernstein, H.D.; Poritz, M.A.; Strub, K.; Hoben, P.J.; Brenner, S.; Walter, P.
Nature 340, 482-486, 1989
A:Title: Model for signal sequence recognition from amino-acid sequence of 54K subunit
A:Reference number: S05198; MUID:89330597; PMID:2502718
A:Accession: S05198
A:Molecule type: mRNA
A:Residues: 1-504 <BERL>
A:Cross-references: EMBL:X16319

Best Local Similarity 32.1%; Pred. No. 1.2e+02;
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;
Qy 44 ELYEGTLGKYNVNEDAKAAATELKSCIDGLOPMHKAELVQLVQLGSDGAG 96
Db 2239 QLLHLTQSEQELEQDA-AARTAMQG--SGLQLPQRSIVQAIARVLGGGCGC 2288
RESULT 15
G86790
fibronectin-binding protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: G86790
R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86790
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-540 <STO>
A:Cross-references: GB:AE005176; PID:gl2724308; PIDN:AAK05425.1; GSPDB:GN00146
A:Experimental source: Strain IL1403
C:Genetics:
A:Gene: yngB
Query Match 12.7%; Score 69; DB 2; Length 540;
Best Local Similarity 30.6%; Pred. No. 25;
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;
Qy 25 DNDEFFM---DFLOTLLVGTPEELYEGTLGKYNVNEDAKAAATELKSCIDGLOP-MHKA 80
Db 294 DNAEIFRQKGLLNTFLNQVPNDKTSVTLENYTNEPIEIALNPALSPVQNAQRVFRYQ 353
Qy 81 LVKLLVQVLGSQ 92
Db 354 KLQAVKFLGEQ 365

Search completed: April 28, 2003, 14:44:35
Job time : 25.732 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:39:06 ; Search time 16.9588 Seconds
(without alignments)
182.172 Million cell updates/sec

Title: US-09-768-826-47
Perfect score: 542
Sequence: 1 MKGRALLVALLFCICRM.....VOVLGSDGACTDYKDDDK 105

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCT05_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	25.2	70	US-09-620-093A-6	Sequence 6, Appli
2	92	17.0	91	US-08-455-896-8	Sequence 8, Appli
3	92	17.0	91	US-08-933-149-8	Sequence 8, Appli
4	92	17.0	91	US-09-082-343-8	Sequence 8, Appli
5	92	17.0	91	US-09-082-253-8	Sequence 8, Appli
6	92	17.0	91	PCT-US96-08235-8	Sequence 8, Appli
7	91	16.8	90	US-08-821-431A-4	Sequence 4, Appli
8	91	16.8	90	US-09-263-810-4	Sequence 4, Appli
9	91	16.8	90	US-08-912-276-15	Sequence 15, Appli
10	91	16.8	90	US-09-583-169-4	Sequence 4, Appli
11	91	16.8	90	US-09-215-818-6	Sequence 6, Appli
12	85.5	15.8	75	US-08-864-357F-3	Sequence 3, Appli
13	79.5	14.7	70	US-08-864-357F-1	Sequence 1, Appli
14	74.5	13.7	69	US-08-864-357F-2	Sequence 2, Appli
15	71	13.1	90	US-08-821-451A-2	Sequence 2, Appli
16	71	13.1	90	US-09-263-810-2	Sequence 2, Appli
17	71	13.1	90	US-09-583-169-2	Sequence 2, Appli
18	69.5	12.8	75	US-08-864-357F-4	Sequence 4, Appli
19	68	12.5	92	US-07-807-529A-2	Sequence 2, Appli
20	68	12.5	92	PCT-US93-02462-2	Sequence 2, Appli
21	68	12.5	94	US-07-662-193-1	Sequence 1, Appli
22	68	12.5	94	US-08-300-928C-2	Sequence 2, Appli
23	68	12.5	94	US-08-430-944D-2	Sequence 2, Appli
24	68	12.5	94	US-08-430-014-2	Sequence 2, Appli
25	68	12.5	94	US-08-431-184-2	Sequence 2, Appli
26	68	12.5	2233	US-08-569-853-1	Sequence 1, Appli
27	68	12.5	2233	US-08-569-853-2	Sequence 2, Appli

28	68	12.5	2233	3	US-08-987-439-1	Sequence 1, Appli
29	66.5	12.3	606	4	US-08-891-298-3	Sequence 3, Appli
30	66.5	12.3	606	4	US-08-653-648A-11	Sequence 11, Appli
31	66	12.2	69	4	US-08-912-276-23	Sequence 23, Appli
32	66	12.2	1139	4	US-09-046-158A-22	Sequence 22, Appli
33	65	12.0	139	4	US-08-559-896B-4	Sequence 4, Appli
34	65	12.0	652	4	US-08-559-896B-2	Sequence 2, Appli
35	63.5	11.7	745	1	US-08-453-472-5	Sequence 5, Appli
36	63.5	11.7	745	1	US-08-038-948-9	Sequence 9, Appli
37	63.5	11.7	745	1	US-08-453-952-5	Sequence 5, Appli
38	63.5	11.7	745	2	US-08-484-993B-43	Sequence 43, Appli
39	63.5	11.7	745	2	US-08-862-903-5	Sequence 5, Appli
40	63.5	11.7	745	2	US-08-484-158B-43	Sequence 43, Appli
41	63.5	11.7	745	2	US-08-484-596A-43	Sequence 43, Appli
42	63.5	11.7	745	2	US-08-480-150A-43	Sequence 43, Appli
43	63.5	11.7	745	3	US-08-458-731-43	Sequence 43, Appli
44	63.5	11.7	745	3	US-08-149-223A-43	Sequence 43, Appli
45	63.5	11.7	1621	1	US-08-242-677-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-620-093A-6
; Sequence 6, Application US/09620093A
; Patent No. 6420337
; GENERAL INFORMATION:
; APPLICANT: Iyer, Subramanian
; APPLICANT: Nguyen, Tay N.
; APPLICANT: Wu, Dauh-Rung
; APPLICANT: Xing, Ruve
; TITLE OF INVENTION: HIGHLY PURIFIED CYTOKINE ACTIVATING FACTOR AND METHODS
; FILE REFERENCE: 3812-7
; CURRENT FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: 60/145,317
; PRIOR FILING DATE: 1999-07-23
; PRIOR APPLICATION NUMBER: 60/197,619
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (54)
; OTHER INFORMATION: xaa = any residue
; NAME/KEY: UNSURE
; LOCATION: (63)
; OTHER INFORMATION: xaa = any residue
US-09-620-093A-6

Query Match 26.2%; Score 142; DB 4; Length 70;
Best Local Similarity 49.2%; Pred. No. 1.2e-10;
Matches 30; Conservative 7; Mismatches 24; Indels 0; Gaps 0;
QY 28 EFFMDFLOLLVGTPEELYEGTIGKYNVNDKAAATELTKSCIDGQPMHKAELVKLLVQ 87
Db 6 EVVPSLIQTLLEGSTEQIYAGPISRYNVDMETSAAALAEKKCIDELPPXHLKALVNLXKQ 65
QY 88 V 88
Db 66 I 66

RESULT 2
US-08-455-896-8
; Sequence 8, Application US/08455896
; Patent No. 5668267

[illegible]

ADDRESSEE: CARELLA, BYRNE, BAIN

QY 1 MKGSRALLVALTLC-----ICRMATGEDNDEFFMDFLQTLVLVGTPEELYEGTLGKYN 54

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

Db 51 PDODMREAGAOLKKLVDTLPKPKPRESIIKLMKIAOS 87

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

```
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,451A
; FILING DATE: March 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,724
; FILING DATE: March 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-821-451A-4

Query Match 16.8%; Score 91; DB 3; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps 2;

QY 1 MKGSRALLVLTLC-----ICRMATGEDNDFMDFLQTLVGTPEELVEGTIGKYN 54
   || || || || || || || || || || || || || || || || || || || ||
Db 1 MKLSVCLLVLTALCCYQANAEFCPALVSELDFFFI-----EPLFKLSLAKFD 50

QY 55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVL 89
   : || : || : || : || : || : || : || : || : || : || : || : ||
Db 51 APPEAVALKGVKRCRDQMSLOKRSLIAEVLVKIL 85

RESULT 8
US-09-263-810-4
; Sequence 4, Application US/09263810
; Patent No. 6174992
; GENERAL INFORMATION:
; APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,810
; FILING DATE:
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,451
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-09-263-810-4

Query Match 16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps 2;

QY 1 MKGSRALLVLTLC-----ICRMATGEDNDFMDFLQTLVGTPEELVEGTIGKYN 54
   || || || || || || || || || || || || || || || || || || || ||
Db 1 MKLSVCLLVLTALCCYQANAEFCPALVSELDFFFI-----EPLFKLSLAKFD 50

QY 55 VNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVL 89
   : || : || : || : || : || : || : || : || : || : || : || : ||
Db 51 APPEAVALKGVKRCRDQMSLOKRSLIAEVLVKIL 85

RESULT 9
US-08-912-276-15
; Sequence 15, Application US/08912276
; Patent No. 6183952
; GENERAL INFORMATION:
; APPLICANT: Billing-Medel, Patricia A.
; APPLICANT: Cohen, Maurice
; APPLICANT: Colpitts, Tracey L.
; APPLICANT: Friedman, Paula N.
; APPLICANT: Gordon, Julian
; APPLICANT: Granados, Edward N.
; APPLICANT: Hodges, Steven C.
; APPLICANT: Klass, Michael R.
; APPLICANT: Kratochvil, Jon D.
; APPLICANT: Roberts-Rapp, Lisa
; APPLICANT: Russell, John C.
; APPLICANT: Stroupe, Steven D.
; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
; TITLE OF INVENTION: FOR DETECTING DISEASES OF THE BREAST
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,276
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L
```

```

Query Match          16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps
0

QY 1 MKGSRALLLVALTLCF-----ICRMATGENDNEFFMDFLQTLTLVGTPEELVEGTLGKYN 54
   || :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MKLUSVCLLVLTALCCYQANAEFCPALVSELLDFFFS-----EPLFKLSLAKFD 50
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 55 VNEDAKAAATELKSCIDGLQPMHKAELVKLLVQVL 89
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 51 APPEAAVAAKLGVRCTDQMSLQKRSLIAEVLVKIL 85
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 11
US-09-215-818-6
: Sequence 6, Application US/09215818A
: Patent No. 6379671
: GENERAL INFORMATION:
: APPLICANT: Colpitts, Tracey
: TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
: TITLE OF INVENTION: DETECTING DISEASES OF THE BREAST
: FILE REFERENCE: 5972 US P2
: CURRENT APPLICATION NUMBER: US/09/215,818A
: CURRENT FILING DATE: 1998-12-18
: EARLIER APPLICATION NUMBER: 08/912,276
: EARLIER FILING DATE: 1997-08-17
: EARLIER APPLICATION NUMBER: 08/697,105
: EARLIER FILING DATE: 1996-08-19
: EARLIER APPLICATION NUMBER: 08/912,149
: EARLIER FILING DATE: 1997-08-15
: EARLIER APPLICATION NUMBER: 08/697,106
: EARLIER FILING DATE: 1996-08-19
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 6
: LENGTH: 90
: TYPE: PRT
: ORGANISM: Homo Sapiens
US-09-215-818-6

Query Match          16.8%; Score 91; DB 4; Length 90;
Best Local Similarity 29.5%; Pred. No. 0.00043;
Matches 28; Conservative 16; Mismatches 35; Indels 16; Gaps
0

QY 1 MKGSRALLLVALTLCF-----ICRMATGENDNEFFMDFLQTLTLVGTPEELVEGTLGKYN 54
   || :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MKLUSVCLLVLTALCCYQANAEFCPALVSELLDFFFS-----EPLFKLSLAKFD 50
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 55 VNEDAKAAATELKSCIDGLQPMHKAELVKLLVQVL 89
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 51 APPEAAVAAKLGVRCTDQMSLQKRSLIAEVLVKIL 85
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 12
US-08-864-357F-3
: Sequence 3, Application US/08864357F
: Patent No. 6255281
: GENERAL INFORMATION:
: APPLICANT: Claragen, Inc. & NIH
: TITLE OF INVENTION: Use of Recombinant Human Uteroglobulin in Treatment
: TITLE OF INVENTION: Fibrotic Conditions
: FILE REFERENCE: 116142/2
: CURRENT APPLICATION NUMBER: US/08/864,357F
: CURRENT FILING DATE: 1997-05-28
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: Patent in version 3.0
: SEQ ID NO 3
: LENGTH: 75
: TYPE: PRT
: ORGANISM: rat
US-08-864-357F-3

Query Match          15.8%; Score 85.5; DB 4; Length 75;

```

```

Db      65 KS 66

RESULT 15
US-08-821-451A-2
; Sequence 2, Application US/08821451A
; Patent No. 6066724
; GENERAL INFORMATION:
; APPLICANT: Jian NI, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,451A
; FILING DATE: March 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,724
; FILING DATE: March 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 90 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
US-08-821-451A-2

Query Match      13.1%; Score 71; DB 3; Length 90;
Best Local Similarity 26.3%; Pred. No. 0.14;
Matches 25; Conservative 14; Mismatches 40; Indels 16; Gaps 2;

QY      1 MKGSRALLVALTLC-----ICRMATGEDNDEFFMDFLQTLTLLVGTPPELYEGTILGKYN 54
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1 MRLSVCLLWVSLALCCYQAHVCPAVASEITVFLFLSDAAVNL-----QVAKLN 50

QY      55 VNEDAKAAMTELSKCIDGLQPMHKAEVLKLLVQVL 89
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      51 PPPEALAAKLEVKHCTDQISFKKRLSLEKVLVEIV 85

Search completed: April 28, 2003, 14:45:28
Job time : 18.9588 secs

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:36:31 ; Search time 20.2887 seconds
(without alignments)
832.774 Million cell updates/sec

Title: US-09-768-826-47_COPY_24_105

Perfect score: 427

Sequence: 1 EDNDEFFMDFLQTLVGTPE.....VQVLGSQDGAGTDYKDDDDK 82

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_cheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	362	84.8	95	4 Q8TD33	Q8td33 homo sapien
2	290	67.9	94	11 Q05702	Q05702 rattus ratt
3	86.5	20.3	478	5 Q62371	Q62371 suberites d
4	76	17.8	2104	16 Q8XM24	Q8xm24 clostridium
5	74.5	17.4	96	11 Q8VD96	Q8vd96 mesocricetu
6	69.5	16.3	2458	2 Q51827	Q51827 pseudomonas
7	69	16.2	540	16 Q9CFY4	Q9cfy4 lactococcus
8	68.5	16.0	2100	3 P87112	P87112 schizosacch
9	68	15.9	1425	4 Q95247	Q95247 homo sapien
10	68	15.9	1483	4 Q95277	Q95277 homo sapien
11	68	15.9	1527	4 Q9UIG0	Q9uig0 homo sapien
12	67.5	15.8	317	10 Q8VXX7	Q8vxx7 arabidopsis
13	67.5	15.8	454	11 Q802R1	Q8qzr1 mus musculus
14	67.5	15.8	485	10 Q9LVG2	Q9lvq2 arabidopsis
15	67.5	15.8	1933	3 Q13450	Q13450 saccharomyc
16	66.5	15.6	338	16 Q8YCG5	Q8ycg5 brucella me

ALIGNMENTS

RESULT 1

ID	Q8TD33	PRELIMINARY;	PRT;	95 AA.
AC	Q8TD33;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DE	Putative ligand binding protein RYD5.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RA	SEQUENCE FROM N.A.			
RP	Bingle C.D.;			
RT	"Human RYD5, a new secretogloblin."			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY026938; AAK08972.1;			
SQ	SEQUENCE 95 AA; 10457 MW; F629AF06C96D2392 CRC64;			

Query Match 84.8%; Score 362; DB 4; Length 95;
Best Local Similarity 98.6%; Pred. No. 1.4e-33;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	EDNDEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLOPMHKAELVK	60
Db	24	EDNDEFFMDFLQTLVGTPEELGKYNVNEDAKAAMTELKSCIDGLOPMHKAELVK	83
QY	61	LLVQVLGSQDGA 72	
Db	84	LLVQVLGSQDGA 95	

RESULT 2

ID	Q05702	PRELIMINARY;	PRT;	94 AA.
AC	Q05702;			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Potential ligand-binding protein.
GN RYD5.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=FISCHER; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=92007724; PubMed=1915264;
RA Dear T.N., Boehm T., Keverne E.B., Rabbitts T.H.;
RT "Novel genes for potential ligand-binding proteins in subregions of
RL the olfactory mucosa.";
RL EMBO J. 10:2813-2819(1991).
DR InterPro: IPR000329; Uterogloblin.
DR SMART: SM00096; UTG; 1
SQ SEQUENCE 94 AA; 10401 MW; 25A4BBB4977E247 CRC64;

Query Match 67.9%; Score 290; DB 11; Length 94;
Best Local Similarity 80.0%; Pred. No. 2.1e-25;
Matches 56; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EDNDEFMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVK 60
Db 23 EDNDEFMEFLQTLVLVGTPEELYEGPLGKYNVNEDAKAALTELKSCIDELQPVHKEQLVK 82
Qy 61 LLVQVLGSQD 70
Db 83 LLVQVLDAQE 92

RESULT 3
ID O62571 PRELIMINARY; PRT; 478 AA.
AC O62571;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Stress-responsive protein kinase (PRKSD).
OS Suberites domuncula (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Hadromerida; Suberitidae; Suberites.
OX NCBI_TaxID=55567;
RN [1]
SEQUENCE FROM N.A.
RP Mueller W.E.G.;
RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RL [2]
SEQUENCE FROM N.A.
RX MEDLINE=96394688; PubMed=8798342;
RA Kruse M., Gamulin V., Cetkovic H., Pancer Z., Mueller I.M.,
RA Mueller W.E.G.;
RT "Molecular evolution of the Metazoan protein kinase C multigene
family.";
RL J. Mol. Evol. 43:374-383(1996).
DR EMBL: Y13101; CAA73555.1; -.
DR HSSP: P24941; IHCL.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 478 AA; 54098 MW; 6B38FF5FD88D6332 CRC64;

Query Match 20.3%; Score 86.5; DB 5; Length 478;
Best Local Similarity 30.6%; Pred. No. 0.22;
Matches 26; Conservative 18; Mismatches 24; Indels 17; Gaps 4;
```

```
Qy 1 EDNDEF---FMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKA 56
Db 240 KDTOKFSNEFSDFIISRLCVKPNPEERMSATALLQHKFIKSAK-PVAVLK----- 286

Qy 57 ELVLLVQVLGSQDAGTDYKDDDD 81
Db 287 ELIADAMRILEEGSGSDEDDND 311

RESULT 4
ID O8XM24 PRELIMINARY; PRT; 2104 AA.
AC O8XM24;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable alpha-N-acetylglucosaminidase.
GN CPE0866.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL: AP003188; BAB80572.1; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000421; FA58_C.
DR InterPro: IPR003961; FN_III.
DR Pfam: PF00754; F5_F8_type_C; 1.
DR Pfam: PF00041; fn3; 1.
DR SMART: SM00060; FN3; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 2104 AA; 5FE345659599BED7 CRC64;

Query Match 17.8%; Score 76; DB 16; Length 2104;
Best Local Similarity 29.0%; Pred. No. 22;
Matches 27; Conservative 10; Mismatches 34; Indels 22; Gaps 3;

Qy 1 EDNDEF---FMDFLQTLVLVGTPEELYECTLGKYNVNEDAKAAMTELKSCIDGLQPMHKA 57
Db 721 KDSAFLYDFADILKQLLANSAQEEYEVCMNAYNNGCKEKFVSGKF-----LE 770

Qy 58 LVKLLVQVLGSQ-----DGAGTDYKDDDD 81
Db 771 LKQLQVRLSTRPEFLIGNWIEDARTMLKSDDD 803

RESULT 5
ID Q8VD96 PRELIMINARY; PRT; 96 AA.
AC Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CC10 protein precursor.
GN CC10.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=LUNG;
RA Gutierrez-Sagal R., Nieto A.;
```


RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99077764; PubMed=9858827;
RA	Peoples R.J., Cisco M.J., Kaplan P., Francke U.;
RT	"Identification of the WBSR9 gene, encoding a novel transcriptional
RT	regulator, in the Williams-Beuren syndrome deletion at 7q11.23.";
RL	Cytogenet. Cell Genet. 82:238-246(1998).
DR	ENBL; AF084479; AAD08675.1; -
DR	HSP; Q92831; IB91
DR	InterPro: IPR001487; Bromodomain.
DR	InterPro: IPR004022; DDT_dom.
DR	InterPro: IPR001965; Znf_PHD.
DR	InterPro: IPR001841; Znf_Ring.
DR	Pfam; PF00439; bromodomain; 1.
DR	Pfam; PF00628; PHD; 1.
DR	PRINTS; PR00503; BROMODOMAIN.
DR	SMART; SM00297; BROMO; 1.
DR	SMART; SM00249; PHD; 1.
DR	SMART; SM00184; RING; 1.
DR	PROSITE; PS50014; BROMODOMAIN_2; 1.
SQ	SEQUENCE 1483 AA; 170903 MW; A47E0365439C5EC2 CRC64;
Query Match	15.9%; Score 68; DB 4; Length 1483;
Best Local Similarity	23.7%; Pred. No. 1.2e+02;
Matches	23; Conservative 20; Mismatches 34; Indels 20; Gaps
QY	2 DNDEFFMDELQTLVGTPEELVEGLTKGYNNVEDAKAAATELK-----SCIDGLQPMHK 55
DB	1046 DGNQLLEFLNRSDLLIEVATRLQGKGLGVETSEFEARVISLEKLDFGCVIALQ----II
QY	56 AEVLKLLVO-----VLGSQDGAGCTDYKDDDK 82
DB	1102 ASVKKFLOGMAPKQRKRKLSQSDSAKTEVDSEKK 1138
RESULT 11	
ID	Q9UGI0 PRELIMINARY; PRT; 1527 AA.
AC	Q9UGI0;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Bromodomain adjacent to zinc finger domain 1B.
GN	BAZ1B.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID	=9606;
NP	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20130112; PubMed=10662543;
RA	Jones M.H., Hamana N., Nezu J., Shimane M.;
RT	"A novel family of bromodomain genes.";
RL	Genomics 63:40-45(2000).
DR	EMBL; AB032253; BAA89210.1; -
DR	HSP; Q92831; I891
DR	InterPro: IPR001487; Bromodomain.
DR	InterPro: IPR004022; DDT_Dom.
DR	InterPro: IPR001965; Znf_PHD.
DR	InterPro: IPR001841; Znf_Ring.
DR	Pfam; PF00439; bromodomain; 1.
DR	Pfam; PF00628; PHD; 1.
DR	PRINTS; PR00503; BROMODOMAIN.
DR	SMART; SM00297; BROMO; 1.
DR	SMART; SM00249; PHD; 1.
DR	SMART; SM00184; RING; 1.
DR	PROSITE; PS50014; BROMODOMAIN_2; 1.
SQ	SEQUENCE 1527 AA; 176033 MW; D11CA368CD55EFIC CRC64;
Query Match	15.9%; Score 68; DB 4; Length 1527;
Best Local Similarity	23.7%; Pred. No. 1.2e+02;
Matches	23; Conservative 20; Mismatches 34; Indels 20; Gaps

```
Oy 2 DNDEFFDFLQTLVGTPEELGKYNVNEDAKAAMTELK-----SCIDGLQPMHK 55
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1046 DGNQELLNFSLDLIEVATRLKGGGLGVETSEFARVISLEKLKDFGECVIALQ----- 1101
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 56 AELVKLLVQ-----VLGSDGAGTDYKDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1102 ASVIKKFLQGPAPKQRRKLOSDSAKTEVDEKK 1138
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q8VXX7 ID Q8VXX7 PRELIMINARY; PRT; 317 AA.
AC Q8VXX7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative APETALA2.
GN AT5G60120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G.S., Bowser L.,
RA Carlinici P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kaniya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene At5g60120 (GI:15238620).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074373; AAL67089.1; -
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR SMART; SM00380; AP2; 1.
DR SEQUENCE 317 AA; 35036 MW; F95F8FF802608C17 CRC64;

Query Match 15.8%; Score 67.5; DB 10; Length 317;
Best Local Similarity 28.8%; Pred. No. 19;
Matches 15; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

Oy 31 NVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQLVGLSDGAGTDYKDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 SVTSNSVVNAEASSCIDGDELCSRTVTKQFEIL---KGGGEEEEDDDE 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
Q8QZR1 ID Q8QZR1 PRELIMINARY; PRT; 454 AA.
AC Q8QZR1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to tyrosine aminotransferase (Hypothetical 50.6 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=LIVER;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; BC024120; AAH24120.1; -
DR EMBL; BC024264; AAH24264.1; -
DR EMBL; BC025934; AAH25934.1; -
KW Hypothetical protein; Aminotransferase; Transferase.
SQ SEQUENCE 454 AA; 50565 MW; CDB6E16CDB5DC7D CRC64;

Query Match 15.8%; Score 67.5; DB 11; Length 454;
Best Local Similarity 33.9%; Pred. No. 30;
Matches 19; Conservative 11; Mismatches 21; Indels 5; Gaps 2;

Oy 11 LQTLVGTPEELGKYNVNED-AKAAMTELKSCIDGLQPMHKAELVLLVQV 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 327 LKSILOQTPOEFYQDTLSFLKSNADLCYGAL-----SAIPGLQPVPSGMYLMVGI 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q9LVG2 ID Q9LVG2 PRELIMINARY; PRT; 485 AA.
AC Q9LVG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE AP2 domain transcription factor-like.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RC MEDLINE=20181125; Pubmed=10718197;
RX Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty P1 and TAC clones.";
RL DNA Res. 7:31-63(2000).
DR EMBL; AB019231; BAA96941.1; -
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR SMART; SM00380; AP2; 1.
DR SEQUENCE 485 AA; 53164 MW; BEFB290E407FC3D5 CRC64;

Query Match 15.8%; Score 67.5; DB 10; Length 485;
Best Local Similarity 28.8%; Pred. No. 33;
Matches 15; Conservative 10; Mismatches 24; Indels 3; Gaps 1;

Oy 31 NVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQLVGLSDGAGTDYKDDDK 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 SVTSNSVVNAEASSCIDGDELCSRTVTKQFEIL---KGGGEEEEDDDE 82
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
O13450 ID O13450 PRELIMINARY; PRT; 1953 AA.
AC O13450;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 'PPF3P'.
GN BNI1 OR PPF3 OR YNL271C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180;
RA Yorihuzi T., Ohsumi Y.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D38411; BAA22512.1; -
DR SGD; S0005215; BNI1.
```

DR InterPro; IPR003104; FH2.
DR Pfam; PF02181; FH2; 1.
DR SMART; SM00498; FH2; 1.
SQ SEQUENCE 1953 AA; 219671 MW; 0C5A9D5280A5858F CRC64;

Query Match 15.8%; Score 67.5; DB 3; Length 1953;
Best Local Similarity 32.9%; Pred. No. 1.9e+02;
Matches 24; Conservative 6; Mismatches 20; Indels 23; Gaps 4;

QY 13 TLLVGTPEELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQVLGSDG- 71
|:| ||| :||| :| | || :| | |||
Db 804 TILNSRPEE-----SFKNNEGTRGMSSLNS-----SEKANIQKVL-----QDGL 843

QY 72 --AGTDYKDDDK 82
| | ||||| |
Db 844 SRAKKDYKDDSKK 856

Search completed: April 28, 2003, 14:43:29
Job time : 23.2887 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:33:15 ; Search time 39.1684 Seconds
(without alignments)
278.964 Million cell updates/sec

Title: us-09-768-826-47_COPY_24_105

Perfect score: 427

Sequence: 1 EDNDEPFMDFLTLVGTPE.....VQVLGSDGAGTDYKDDDK 82

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	427	100.0	105	22	Human secreted pro-
2	427	100.0	105	23	Human albumin fusi
3	362	84.8	95	23	Human secreted pro
4	142	33.3	70	22	Cytokine activatin
5	85.5	20.0	75	20	Rat uteroglobin-li
6	85.5	20.0	75	21	Rat uteroglobin-li
7	81	19.0	96	18	hPSTI, SSI chimeric
8	79.5	18.6	70	15	Lymphocytoma proli
9	79.5	18.6	70	18	Target peptide fro
10	79.5	18.6	70	20	Human uteroglobin-

11	79.5	18.6	70	21	Human uteroglobin.
12	79.5	18.6	91	22	Human Clara cell p
13	76.5	17.9	78	18	hPSTI,OMTKY3 chime
14	76.5	17.9	78	18	hPSTI,SSImutant ch
15	76.5	17.9	102	22	Bovine mammary tis
16	76.5	17.9	452	22	Novel human diagno
17	74.5	17.4	69	20	Rabbit uteroglobin
18	74.5	17.4	69	21	Rabbit uteroglobin
19	73	17.1	936	22	Murine mature chor
20	73	17.1	962	22	Murine chordin (CH
21	69.5	16.3	75	20	Mouse uteroglobin-
22	69.5	16.3	75	21	Murine uteroglobin
23	69.5	16.3	119	21	Lung cancer associ
24	69.5	16.3	187	23	Mature human LP-48
25	69	16.2	47	19	Ad5 fiber knob dom
26	68	15.9	540	23	Lactococcus lactis
27	68	15.9	1527	19	Human BAZ1-beta pr
28	68	15.9	1531	19	Human BAZ1-beta pr
29	67	15.7	167	23	Protein rCD44HS.
30	67	15.7	423	22	Human polypeptide
31	67	15.7	465	22	Human polypeptide
32	66	15.5	95	21	Borna disease viru
33	66	15.5	1139	19	Human JAK2 protein
34	65.5	15.3	457	22	Drosophila melanog
35	65	15.2	139	18	Plasmodium falcipa
36	65	15.2	323	22	Murine mature chor
37	65	15.2	345	22	Murine chordin-lik
38	65	15.2	652	18	Plasmodium falcipa
39	65	15.2	842	23	Herbicidally activ
40	64.5	15.1	92	23	Human ORF1596 prot
41	64.5	15.1	1594	23	Modified human/mou
42	64	15.0	24	22	Flag epitope tag f
43	64	15.0	2086	22	Staphylococcus aur
44	64	15.0	5795	22	Staphylococcus aur
45	64	15.0	6281	22	Staphylococcus aur

ALIGNMENTS

RESULT 1

AAAB70079

ID AAB70079 standard; Protein; 105 AA.

XX AAB70079;

XX 14-MAY-2001 (first entry)

DT Human secreted protein #18.

DE Human; secreted protein; immunomodulatory; antisclerotic;

XX dermatological; immunosuppressive; antiinflammatory; anti-HIV;

XX immunostimulant; cytostatic; cardiant; vascular; angiogenic;

KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;

KW antiallergics; antiparkinsonian; antimicrobial; vulnary; gene therapy;

KW immune disorder; hyperproliferative; cardiovascular; angiogenic;

KW neurological; infection.

XX Homo sapiens.

OS WO200112776-A2.

PN 22-FEB-2001.

PD 15-AUG-2000; 2000WO-US22350.

XX 16-AUG-1999; 99US-0148759.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;

XX WPI; 2001-244245/25.

DR N-PSDB; AAF76860.
XX
XX Nucleic acids encoding 18 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
XX Claim 11; Page 373; 380pp; English.
XX
XX The present sequence is one of 18 novel human secreted proteins. The
CC nucleic acids encoding the proteins and the proteins themselves may be
CC used in the prevention, diagnosis and treatment of diseases including
CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
CC and human immunodeficiency virus (HIV) infections), hyperproliferative
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC neovascularisation and diabetic retinopathy), neurological disorders
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC infectious diseases and/or for promoting wound healing, regeneration
CC and/or chemotaxis. The nucleic acid molecules may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples. The polypeptides may also be
CC used as antigens in the production of antibodies and in assays to
CC identify modulators of protein expression and activity.
XX
XX Sequence 105 AA;
SQ
Query Match 100.0%; Score 427; DB 22; Length 105;
Best Local Similarity 100.0%; Pred. No. 5.7e-45;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 60
DB 24 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 83
QY 61 LLVQVLGSDGAGTDYKDDDDK 82
DB 84 LLVQVLGSDGAGTDYKDDDDK 105
RESULT 2
ABG65513
ID ABG65513 standard; Protein; 105 AA.
XX
XX AC ABG65513;
XX
XX 27-AUG-2002 (first entry)
XX
XX Human albumin fusion protein #2188.
XX
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antiulcer;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO20017137-A1.
XX
XX 18-OCT-2001.
PD
XX 12-APR-2001; 2001WO-US11988.
PF
XX 12-APR-2000; 2000US-229358P.
PR
XX 25-APR-2000; 2000US-199384P.
PR
XX 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA

XX Rosen CA, Haseltine WA;
PI WPI; 2002-010886/01.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
PT albumin fused to a therapeutic protein -
PT
XX Claim 1; Page 2082; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or
CC disorder that may be modulated by therapeutic protein X. The albumin
CC extends the shelf-life of protein X, and may increase its biological
CC in vitro/in vivo activity. The protein is useful for treating and
CC diagnosing disorders such as cancer, reproductive disorders, digestive
CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
CC (e.g. diabetes), haematopoietic disorders, neural disorders
CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
CC encephalomyelitis, meningitis, schizophrenia), and connective disorders
CC (e.g. osteoporosis, arthritis). ABG6326-ABG65518 represent albumin
CC fusion proteins of the invention.
XX
XX Sequence 105 AA;
SQ
Query Match 100.0%; Score 427; DB 23; Length 105;
Best Local Similarity 100.0%; Pred. No. 5.7e-45;
Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 60
DB 24 EDNDEFFMDFLQTLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLQPMHKAELVK 83
QY 61 LLVQVLGSDGAGTDYKDDDDK 82
DB 84 LLVQVLGSDGAGTDYKDDDDK 105
RESULT 3
AAO21685
ID AAO21685 standard; Protein; 95 AA.
XX
XX AC AAO21685;
XX
XX 05-SEP-2002 (first entry)
XX
XX Human secreted protein SEQ ID No 27.
DE
XX
XX Antiarteriosclerotic; cytostatic; HIV; antiallergic; antianaemic;
KW antiasthmatic; cardiant; vasotropic; neuroprotective; nootropic; SBCP;
KW anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
KW immunosuppressive; human secreted protein; cell proliferative disorder;
KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
KW ischaemic heart disease; congestive heart failure; neurological disorder;
KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
KW Parkinson's disease; epilepsy; stroke; knockin humanised animal;
KW transgenic animal; gene therapy.
XX
XX Homo sapiens.
OS
XX WO200238602-A2.
XX
XX 16-MAY-2002.
PD
XX 08-NOV-2001; 2001WO-US47420.
PF
XX 08-NOV-2000; 2000US-247505P.
PR
XX 09-NOV-2000; 2000US-248642P.
PR
XX 16-NOV-2000; 2000US-249824P.
PR
XX 21-NOV-2000; 2000US-252824P.
PR

PR 08-DEC-2000; 2000US-254305P.
XX 18-DEC-2000; 2000US-256448P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Wallia NK;
PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;
PI Ding L, Hafalia AJA, Tang YT, Bandman O, Warren BA, Honchell CD;
PI Lu DAM, Thangavelu K, Lee S, Xu Y, Yang J, Lal PG, Tran B;
PI Ison CH, Duggan BM, Sapperstein SK;
XX WPI; 2002-519296/55.
DR N-PSDB; AAL39646.
XX Human secreted proteins and polynucleotides for diagnosing, treating or
XX preventing disorders of cell proliferative, cardiovascular,
XX developmental, neurological and autoimmune/inflammatory disorders -
XX Claim 1; Page 168; 229pp; English.
XX The invention relates to an isolated human secreted protein (SECP)
XX polypeptide from 63 fully defined human sequences given in the
XX specification. The polypeptide is useful for the diagnosing/treating of a
XX disease with decreased/overexpression of SECP. Examples of disorders
XX associated with abnormal expression of SECP include a cell proliferative
XX disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
XX disorder, AIDS, allergies, anaemia, asthma; cardiovascular disease e.g.
XX congestive heart failure, ischaemic heart disease; developmental disorder
XX e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
XX Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
XX The SECP polynucleotide and polypeptide are further useful for analysing
XX the proteome of a tissue or a cell type. The polynucleotide is useful for
XX creating knockin humanised animals (pigs) or transgenic animals (mice or
XX rats) to model human disease, and for somatic or germline gene therapy,
XX and further for generating hybridisation probes useful in mapping the
XX naturally occurring genomic sequence. This sequence represents a human
XX secreted protein of the invention.
XX Sequence 95 AA:
SQ
Query Match 84.8%; Score 362; DB 23; Length 95;
Best Local Similarity 98.6%; Pred. No. 5.5e-37;
Matches 71; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EDNDEFFMDFLQTLVGNPEELVGTGLGKYNVNEDAKAAMTELKSCIDGLOPMHKAELVK 60
Db 24 EDNDEFFMDFLQTLVGNPEELVGTGLGKYNVNEDAKAAMTELKSCIDGLOPMHKAELVK 83
Qy 61 LLVQVLGSQDGA 72
Db 84 LLVQVLGSQDGA 95
RESULT 4
AAB72211
ID AAB72211 standard; protein; 70 AA.
XX AAB72211;
XX AAB72211;
DT 04-MAY-2001 (first entry)
XX Cytokine activating factor CAFb amino acid sequence.
XX Cytokine activating factor; CAF; tumour necrosis factor alpha; TNF-alpha;
KW interleukin-beta; IL-beta; transforming growth factor beta; TGF-beta;
KW immune response; cancer; sepsis; septic shock; bacterial infection;
KW chicken.
XX Gallus sp.
OS
XX WO200107472-A1.
PN
XX 01-FEB-2001.
PD

XX 19-JUL-2000; 2000WO-US19643.
XX 23-JUL-1999; 99US-0145317.
PR 14-APR-2000; 2000US-0197619.
XX (DCVD-) DCV INC.
XX Iyer S, Nguyen TN, Wu D, Xing R;
PI WPI; 2001-123262/13.
XX New cytokine activating factor, useful for regulating an immune
XX response and treating cancer and sepsis in an animal -
XX Claim 1; Page 61; 96pp; English.
XX This invention relates to a cytokine activating factor (CAF) protein
XX which upregulates the expression of tumour necrosis factor alpha
XX (TNF-alpha), interleukin-beta (IL-beta) or downregulates expression of
XX transforming growth factor beta (TGF-beta). A composition, preferably a
XX food product, containing a CAF protein is useful for regulating an immune
XX response and treating cancer in an animal. It is also useful for treating
XX or preventing sepsis and/or septic shock and bacterial infection in an
XX animal. The present sequence represents the full length CAFb protein.
XX Sequence 70 AA:
SQ
Query Match 33.3%; Score 142; DB 22; Length 70;
Best Local Similarity 49.2%; Pred. No. 6.2e-10;
Matches 30; Conservative 7; Mismatches 24; Indels 0; Gaps 0;
Qy 5 EFFEFDLQTLVGNPEELVGTGLGKYNVNEDAKAAMTELKSCIDGLOPMHKAELVKLLVQ 64
Db 6 EWFSLQTLLEGSTIEQYAGPISRYNVDEMTSAALAEIKKICIDELPPXHLKALVNLXKQ 65
Qy 65 V 65
Db 66 I 66
RESULT 5
AAW87858
ID AAW87858 standard; protein; 75 AA.
XX AAW87858;
AC AAW87858;
XX 12-MAR-1999 (first entry)
DT Rat uteroglobin-like protein CC10.
XX Uteroglobin; phospholipase A2; fibronectin; inflammation; asthma;
KW cystic fibrosis; premature labour; infertility; rheumatoid arthritis;
KW type I diabetes; nephropathy; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; pancreatitis; peritonitis; allergy;
KW multiple organ failure; adult respiratory distress syndrome;
KW acute renal failure; organ transplant rejection; autoimmune uveitis;
KW corneal transplant surgery; neonatal RDS; cytomegalovirus retinitis;
KW pneumonia; cystitis; schistosomiasis; vaginal candidiasis; fibrosis;
KW neonatal broncho-pulmonary dysplasia; haemodialysis; glomerulopathy;
KW artificial insemination.
XX Rattus sp.
OS
XX WO9853846-A1.
PN
XX 03-DEC-1998.
PD 28-MAY-1998; 98WO-US11026.
XX 28-MAY-1997; 97US-0864357.
PR (CLAR-) CLARAGEN INC.
XX

PR 25-OCT-1995; 95US-0548186.
XX (ARRI-) ARRIS PHARM CORP.
XX Estell DA, Fisher J, Hartman C, McGrath M, Schmidt B;
XX WPI: 1997-259023/23.
DR N-PSDB; AAT84524.
XX
PT Modified subtilisin inhibitor - having altered reactive site
PT resulting in decreased immunogenicity, used in detergent
PT composition(s)
XX
XX Example 2; Fig 9A; 69pp; English.
PS
CC This polypeptide comprises human pancreatic secretory trypsin
CC inhibitor (PSII), modified to include a Streptomyces subtilisin
CC inhibitor (SSI) reactive site. A phoA signal peptide facilitates
CC expression in E. coli. The polypeptide is encoded by a DNA
CC construct (AAT84524) produced by PCR amplification of overlapping
CC synthetic oligonucleotides. Novel modified subtilisin inhibitors
CC are claimed in which a human standard mechanism inhibitor, such as
CC PSII, has an altered reactive site so that its dissociation
CC constant for inhibition of subtilisin is reduced by at least a
CC factor of 100. Also claimed are: a nucleic acid encoding the
CC subtilisin inhibitor; expression vectors; and transformed host
CC cells. The novel subtilisin inhibitors combine the low
CC allergenicity of human standard mechanism inhibitors, which are
CC not specific for subtilisin, and the high affinity of non-human
CC subtilisin inhibitors such as SSI. They form a complex with
CC subtilisin, so as to avoid problems of allergenicity in detergent
CC formulations, and can also be used in affinity purification and
CC (diagnostic) quantification of subtilisin.
XX
SQ Sequence 96 AA;

Query Match 19.0%; Score 81; DB 18; Length 96;
Best Local Similarity 28.1%; Pred. No. 0.033;
Matches 25; Conservative 8; Mismatches 30; Indels 26; Gaps 3;

Qy 15 LVGTPPEELVGTGKYNV-----NEDAKAAM-----TELKSCIDGLQPM 53
Db 13 LLFTPSQAAGNTGADSLGREAKYCNEDVMCPMYPVCGTGTGNTYPCVLCFEG 53
Qy 54 HKAEVLKLVGVLSQDGAGDYKDDDDK 82
Db 69 -RKRTSILIOKSGPCGGGGDYKDDDDK 96

RESULT 8
AAR55778
ID AAR55778 standard; peptide; 70 AA.
AC AAR55778;
XX
XX 21-FEB-1995 (first entry)
XX
XX Lymphocytoma proliferation activating peptide (LPAP).
XX
KW Lymphocytoma proliferation activating peptide; LPAP; tumour;
KW degenerative disease; lymphocyte; mesenchyme tissue;
KW epidermal injury; cardiovascular disease; respiratory disease;
KW urogenital disease; nervous system; immune system; antibody;
KW diagnosis.
XX
OS Homo sapiens.
XX
XX key Location/Qualifiers
XX Disulfide-bond 3
XX
XX /note= "When the peptide exists in dimeric form the
XX cysteine at position 3 forms a disulfide bond
XX with the cysteine at position 69 on the
XX antiparallel peptide strand."

FT Disulfide-bond 69
FT /note= "When the peptide exists in dimeric form the
FT cysteine at position 69 forms a disulfide bond
FT with the cysteine at position 3 on the
FT antiparallel peptide strand."
XX
XX DE4244565-A.
XX
XX 07-JUL-1994.
XX
XX 30-DEC-1992; 92DE-4244565.
XX
XX 30-DEC-1992; 92DE-4244565.
XX (FORS/) FORSMANN W.
XX Adermann K, Aoki A, Forssmann W, Hock D, Meyer M;
XX Raida M, Schulz-knappe;
XX WPI: 1994-219093/27.
XX
XX Lymphocytoma proliferation activating peptide structure and
XX peptide - its prodn and its use for the treatment of, eg tumours,
XX hypoplasia, skin-, cardiovascular-, respiratory- and
XX urogenital-diseases
XX
XX Claim 1; Page 16; 38pp; German.
XX
XX Lymphocytoma proliferation activating peptide (LPAP) is used as a
XX medicament for the treatment of disturbances in the cell growth of
XX normal and tumour cells; degenerative diseases of humans e.g. of
XX lymphocytes and other mesenchyme tissues; epidermal injuries;
XX cardiovascular-, respiratory- and urogenital diseases; the nerve
XX system and the immune system. Antibodies directed against LPAP can
XX also be used in diagnostic assays.
XX
SQ Sequence 70 AA;

Query Match 18.6%; Score 79.5; DB 15; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;

Qy 7 FMDFLQTLVGTPEELVGTGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQYL 66
Db 6 FQRTVETLLMDTPSS-YEAMELFSPDQDMREAGAGLKKLVDTLPQKPRESIKLMKXIA 64
Qy 67 GS 68
Db 65 QS 66

RESULT 9
AAW11628
ID AAW11628 standard; peptide; 70 AA.
XX
XX AAW11628;
XX
XX 07-OCT-1997 (first entry)
XX
XX Target peptide from an ubiquitin fusion protein.
XX
XX Recombinant peptide; host cell; expression vector; hormone;
XX antimicrobial; epitope; ACTH; VIP; GHRH; CGRP; amylin; osteocalcin;
XX insulin; natriuretic peptide.
XX
XX Synthetic.
XX
XX WO9701627-A1.
XX
XX 16-JAN-1997.
XX
XX 27-JUN-1996; 96WO-US10858.
XX

```

PR 03-NOV-1995; 95US-0553006.
PR 27-JUN-1995; 95US-0000822.
XX (IGEN-) IGEN INC.
PA (YOST/) YOST P B.
XX
PI Lohnas GL, Pilon AL, Roberts SF, Yost PB;
XX WPI: 1997-100203/09.
XX
XX High level expression of recombinant peptide(s) - using host cells
PT comprising an expression vector encoding a ubiquitin-peptide fusion
PT protein
XX
XX Claim 40; Page 56; 72pp; English.
XX
XX A fermentation method has been discovered for producing a target peptide
CC from an ubiquitin fusion protein. The method involves inducing the
CC synthesis of the fusion protein by a transformant contained in a
CC fermentation culture; culturing the transformant under conditions
CC suitable for intracellular production and accumulation of the induced
CC fusion protein; recovering the fusion protein in at least 50% specific
CC yield; hydrolysing the induced fusion protein using an ubiquitin
CC specific hydrolase to release the target peptide; and recovering the
CC target peptide. The present sequence represents a specifically claimed
CC example of a target peptide. The method can be used to produce e.g.
CC peptide hormones, antimicrobial peptides, peptide epitope fragments,
CC ACTH, VIP, GRH, CGRP, amylin, osteocalcin, insulin, and natriuretic
CC peptides. The system can provide recombinant fusion proteins in high
CC specific yields of over 50% of total cellular protein.
XX
XX Sequence 70 AA;
SQ
Query Match 18.6%; Score 79.5; DB 18; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
Qy 7 FMDFLQTLVGTPELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQL 66
Db 6 FORVIETLLMDTPSS-YEAAANELFSPDQDMREAGAKLKLVDLTLPQKPRESIKLMEXIA 64
Qy 67 GS 68
Db 65 QS 66
RESULT 10
AAW87569
ID AAW87569 standard; protein; 70 AA.
XX
AC AAW87569;
XX
XX 12-MAR-1999 (first entry)
XX
DE Human uteroglobin-like protein CC10.
XX
XX Uteroglobin; phospholipase A2; fibronectin; inflammation; asthma;
KW cystic fibrosis; premature labour; infertility; rheumatoid arthritis;
KW type I diabetes; nephropathy; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; pancreatitis; peritonitis; allergy;
KW multiple organ failure; adult respiratory distress syndrome;
KW acute renal failure; organ transplant rejection; autoimmune uveitis;
KW corneal transplant surgery; neonatal RDS; cytomegalovirus retinitis;
KW pneumonia; cystitis; schistosomiasis; vaginal candidiasis; fibrosis;
KW neonatal broncho-pulmonary dysplasia; haemodialysis; glomerulopathy;
KW artificial insemination.
XX
OS Homo sapiens.
XX
XX WO9853846-A1.
XX
XX 03-DEC-1998.
XX

```

```

PF 28-MAY-1998; 98WO-US11026.
XX
PR 28-MAY-1997; 97US-0864357.
XX
PA (CLAR-) CLARAGEN INC.
PA (USSH ) US NAT INST OF HEALTH.
XX
PI Mukherjee AB, Pilon AL, Zhang Z;
XX WPI: 1999-059777/05.
XX
XX Treating and preventing inflammation and fibrosis with human
PT uteroglobin - which inhibits phospholipase A2 and binds to
PT fibronectin, for clinical or cosmetic use, e.g. in respiratory
PT distress syndrome
XX
XX Disclosure; Fig 1; 59pp; English.
XX
XX The present sequence represents a uteroglobin-like protein. Recombinant
CC human uteroglobin inhibits phospholipase A2 (PLA2), and binds to
CC fibronectin. Inhibition of PLA2 is used to treat or prevent a wide range
CC of systemic and ocular inflammations, asthma, cystic fibrosis, premature
CC labour, infertility, rheumatoid arthritis, type I diabetes, nephropathy,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC pancreatitis, peritonitis, allergy, multiple organ failure, adult
CC respiratory distress syndrome (RDS), acute renal failure, inflammation
CC secondary to infection or surgery, and organ transplant rejection. Some
CC specified applications are in autoimmune uveitis, corneal transplant
CC surgery, neonatal and adult RDS, cytomegalovirus retinitis, pneumonia,
CC cystitis, schistosomiasis and vaginal candidiasis. Fibrotic conditions
CC that can be treated are pulmonary, renal and vascular fibrosis.
CC Uteroglobin may be administered to correct deficiency in endogenous
CC uteroglobin, e.g. in neonatal broncho-pulmonary dysplasia, complications
CC of haemodialysis and inherited glomerulopathy. Uteroglobin can also be
CC used to increase the rate of artificial insemination, in humans or
CC animals, by treatment of sperm, fertilised eggs or embryos before
CC transfer to the uterus.
XX
XX Sequence 70 AA;
SQ
Query Match 18.6%; Score 79.5; DB 20; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
Qy 7 FMDFLQTLVGTPELYEGTLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVKLLVQL 66
Db 6 FORVIETLLMDTPSS-YEAAANELFSPDQDMREAGAKLKLVDLTLPQKPRESIKLMEXIA 64
Qy 67 GS 68
Db 65 QS 66
RESULT 11
AAB32128
ID AAB32128 standard; protein; 70 AA.
XX
AC AAB32128;
XX
XX 14-FEB-2001 (first entry)
XX
DE Human uteroglobin.
XX
XX Human; uteroglobin; immunoglobulin A mediated disease; IgA nephropathy;
KW autoimmune disorder; pulmonary inflammation; Wegener's granulomatosis;
KW Goodpasture's disease; diabetic glomerulosclerosis.
XX
OS Homo sapiens.
XX
XX WO200062795-A2.
XX
XX 26-OCT-2000.
XX

```

PF 13-APR-2000; 2000WO-US09979.
XX
PR 21-APR-1999; 99US-0130434.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Mukherjee AB, Zheng F, Zhang Z;
XX
XX WPI; 2000-687100/67.
XX
XX Use of a composition comprising uteroglobin (or a fragment, derivative,
PT mimetic or variant), for inhibiting or treating an immunoglobulin-A
PT mediated autoimmune disorders, e.g. diabetic glomerulosclerosis and
PT pulmonary inflammation -
XX
XX Example 9; Fig 1; 60pp; English.
XX
XX The present invention describes the use of uteroglobin in the diagnosis
CC and prevention of IgA mediated diseases, such as IgA nephropathy,
CC Wegener's granulomatosis, Goodpasture's disease and diabetic
CC glomerulosclerosis. This is possible as uteroglobin binds to fibronectin,
CC preventing the complexing of fibronectin with IgA and the deposition of
CC immune complexes in the kidney.
XX
XX Sequence 70 AA:
SQ
Query Match 18.6%; Score 79.5; DB 21; Length 70;
Best Local Similarity 27.4%; Pred. No. 0.034;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
QY 7 FMDFLQTLVLTGTPPELYEGTGLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQVL 66
DB 6 FQVRIETLLMDTPSS-YEAAAMELSPDQDMREGAQLKLVDTLPQKPRESIIKLMKIA 64
QY 67 GS 68
DB 65 QS 66
RESULT 12
AAB64165
ID AAB64165 standard; peptide: 91 AA.
XX
AC AAB64165;
XX
DT 27-MAR-2001 (first entry)
XX
XX Human Clara cell protein-derived antiviral peptide, SEQ ID NO:118.
XX
XX Human Clara cell protein-derived peptide; CC10; antiviral;
KW immune potentiator; anti-infective; therapy; prophylaxis;
KW Clara cell receptor-independent; inhibit expression inducer.
KW pro-apoptotic; anticancer; non-immunogenic; retrovirus; togavirus;
KW hepatitis virus; cytomegalovirus; hantavirus; HIV; human herpes virus;
KW septic shock syndrome; cachexia; tumour; reverse transcriptase inhibitor;
KW p24 inhibitor.
XX
XX Homo sapiens.
OS
XX
XX WO200072868-A2.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 01-JUN-2000; 2000WO-IB00895.
PF
XX
XX 01-JUN-1999; 99US-0137227.
PR
XX 03-MAR-2000; 2000IE-0000170.
PR
XX 13-APR-2000; 2000IE-0000303.
PR
XX 18-MAY-2000; 2000IE-0000385.
PR
XX 18-MAY-2000; 2000IE-0000386.
PR
XX (PREN/) PRENDERGAST P T.
PA

XX Prendergast PT;
PI
XX WPI; 2001-146767/15.
DR
XX
XX Novel pharmaceutical formulation comprising a compound that comprises a
PT peptide isolated from Clara cell protein useful for treating autoimmune
PT disorders, cancers, viral infections, septic shock syndrome and
PT cachexia -
XX
XX Claim 1; Page 129; 139pp; English.
PS
XX
XX The invention relates to pharmaceutical compositions comprising
CC therapeutic peptides corresponding to human Clara cell protein
CC fragments or conservative variants thereof (AAB64050-B64178). The
CC peptides of the invention, unlike native Clara cell protein, have
CC potent antiviral activity, being able to block viral replication via
CC the inhibition of reverse transcriptase and p24 activities. The
CC Clara cell-derived proteins of the invention also have a broad immune
CC potentiator effect and anti-infective (antibacterial and antifungal, as
CC well as antiviral) activity. The peptides of the invention may also
CC induce the synthesis of inhibin, a secreted protein with tumour
CC suppressor activity, and may also upregulate the expression of
CC pro-apoptotic genes. The peptides may additionally mimic the effects of
CC full-term pregnancy with respect to long-term protection against
CC chemical carcinogens and lifetime reduction in breast cancer risk,
CC and also inhibit prostaglandin-induced pre-term delivery. The
CC peptides induce functional maturation of cortical thymocytes, and act as
CC a factor regulating antigen-independent differentiation of T-lymphocytes
CC during pregnancy, thus enhancing the immune system in embryogenesis.
CC The peptides of the invention are non-immunogenic, and function without
CC the requirement for Clara cell receptors. The peptides are used
CC prophylactically and therapeutically against viral infections,
CC specifically infections caused by retroviruses, togaviruses,
CC flaviviruses, rubiviruses, pestiviruses and hantaviruses, and
CC particularly infections by HIV, human T-lymphocyte virus 1 (HTLV-1),
CC HTLV-3, Kaposi's Sarcoma-associated herpes virus, human herpes virus 6
CC (HHV-6), HHV-8, viruses of the genus Molluscipoxvirus, hepatitis A, B,
CC or C virus, or cytomegalovirus. The peptides can also be used to reduce
CC the likelihood of septic shock syndrome in a patient about to undergo
CC surgery, and can be used to treat this condition in a post-operative
CC patient. They can also reduce the likelihood of cachexia in a patient
CC about to undergo surgery. The peptides can be used to treat autoimmune
CC conditions such as multiple sclerosis, systemic lupus erythematosus,
CC myasthenia gravis, rheumatoid arthritis, Sjogren's disease, ulcerative
CC colitis, and diabetes or to ameliorate immune damage and/or disease
CC progression. The peptides can also be used to treat cancer, or to slow
CC progression or development of cancers such as breast cancer, colon
CC cancer, leukaemia, brain cancer, lung cancer and melanoma. The
CC present sequence represents a human Clara cell protein-derived peptide
XX of the invention.
XX Sequence 91 AA:
SQ
Query Match 18.6%; Score 79.5; DB 22; Length 91;
Best Local Similarity 27.4%; Pred. No. 0.048;
Matches 17; Conservative 17; Mismatches 27; Indels 1; Gaps 1;
QY 7 FMDFLQTLVLTGTPPELYEGTGLGKYNVNEDAKAAMTELKSCIDGLQPMHKAELVLLVQVL 66
DB 27 FQVRIETLLMDTPSS-YEAAAMELSPDQDMREGAQLKLVDTLPQKPRESIIKLMKIA 85
QY 67 GS 68
DB 86 QS 87
RESULT 13
AAW26363
ID AAW26363 standard; Protein; 78 AA.
XX
XX AAW26363;
XX


```
XX SQ Sequence 78 AA;
Query Match 17.9%; Score 76.5; DB 18; Length 78;
Best Local Similarity 35.4%; Pred. No. 0.092;
Matches 17; Conservative 5; Mismatches 21; Indels 5; Gaps 1;

Qy 35 DAKAAMTELKSCIDGLOPMHKAELVKLVQVLGSDGAGTDYKDDDDK 82
| | | | | : | | | | | : | | | | |
Db 36 DGNYPNECVLCFEG-----RKROTSILIOKSGCGGGSDYKDDDDK 78

RESULT 15
AAB87666
ID AAB87666 standard; protein; 102 AA.
AC AAB87666;
XX 15-MAY-2001 (first entry)
DT Bovine mammary tissue derived protein #57.
DE Bovine; mammary gland; cancer; tumour; angiogenesis.
KW Bos taurus.
OS WO200114553-A1.
PN 01-MAR-2001.
PD 23-AUG-2000; 2000WO-N200166.
PF 23-AUG-1999; 99US-0150330.
PR (GENE-) GENESIS RES & DEV CORP LTD.
PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.
XX Havukkala IJ, Gleen M, Grigor MR, Molenaar AJ;
XX WPI; 2001-226619/23.
XX New polypeptides and polynucleotides encoding the polypeptides, which
PT are expressed in bovine mammary gland tissue, useful for stimulating
PT mammary gland growth or function, or inducing differentiation of milk
PT producing cells
XX Claim 11; Page 83; 97pp; English.
XX The present invention relates to proteins derived from bovine
CC mammary gland cells. The invention is useful for stimulating
CC bovine mammary gland cell growth and function, inhibiting the
CC growth of various mammary gland cancer cells, inhibiting
CC angiogenesis and vascularization of tumours, or modulating
CC the growth of blood vessels in a mammal.
XX SQ Sequence 102 AA;
Query Match 17.9%; Score 76.5; DB 22; Length 102;
Best Local Similarity 30.8%; Pred. No. 0.13;
Matches 20; Conservative 10; Mismatches 26; Indels 9; Gaps 1;

Qy 2 DNDEFFMDFLTLLVGTPEELYEGTLGKYNVNEAKAAMTELKSCIDGLOPMHKAELVKL 61
| | | | | : | | | | | : | | | | |
Db 30 DLTEFFY-----FPDLLYRLSLAKYNAPPEAFAAKMEVKQCTDRFSVKNRLIITNI 80

Qy 62 LVQVL 66
| : |
Db 81 LGKIL 85
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 14:36:31 ; Search time 25.732 Seconds
(without alignments)
832.774 Million cell updates/sec

Title: US-09-768-826-47_COPY_2_105
Perfect score: 537

Sequence: 1 KGRALLLVALTLCICRMA.....VOVLGSDGAGTDYKDDDK 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	87.9	95	4 Q8TD33	Q8td33 homo sapien
2	344.5	64.2	94	11 Q05702	Q05702 rattus ratt
3	90.5	16.9	96	11 Q8VD96	Q8vd96 mesocricetu
4	87.5	16.3	478	5 Q62571	Q62571 suberites d
5	76.5	14.2	504	11 Q9D008	Q9d008 mus musculu
6	76	14.2	2104	16 Q8XM24	Q8xm24 clostridium
7	74	13.8	457	5 Q9VAB8	Q9vab8 drosophila
8	73	13.6	90	6 Q9GK57	Q9gk57 oryctolagus
9	73	13.6	609	3 Q9UT83	Q9ut83 schizosacch
10	72.5	13.5	500	11 Q99J29	Q99j29 mus musculu
11	70.5	13.1	3071	12 Q8QXL2	Q8qxl2 sorghum mos
12	70.5	13.1	3071	12 Q8QXL1	Q8qxl1 sorghum mos
13	69.5	12.9	643	10 Q9FW93	Q9fw93 oryza sativ
14	69.5	12.9	2458	2 Q51827	Q51827 pseudomonas
15	69	12.8	540	16 Q9CFY4	Q9cfy4 lactococcus
16	68.5	12.8	2100	3 P87112	P87112 schizosacch

17	68	12.7	1425	4 Q95247	Q95247 homo sapien
18	68	12.7	1483	4 Q95277	Q95277 homo sapien
19	68	12.7	1527	4 Q9UTG0	Q9utg0 homo sapien
20	68	12.7	2258	12 Q81081	Q81081 human parai
21	67.5	12.6	249	17 Q8TY89	Q8ty89 methanopyru
22	67.5	12.6	299	5 Q19742	Q19742 caenorhabdi
23	67.5	12.6	317	10 Q8VXX7	Q8vxx7 arabidopsis
24	67.5	12.6	454	11 Q8QZK1	Q8qzr1 mus musculu
25	67.5	12.6	485	10 Q9LVG2	Q9lvq2 arabidopsis
26	67.5	12.6	1953	3 Q13450	Q13450 saccharomyc
27	67	12.5	511	10 Q80740	Q80740 arabidopsis
28	66.5	12.4	322	16 Q9HVL5	Q9hvl5 pseudomonas
29	66.5	12.4	358	16 Q8YCG5	Q8yrg5 bruceella me
30	66	12.3	151	10 Q8RYJ9	Q8ryj9 oryza sativ
31	66	12.3	189	2 Q85808	Q85808 borrelia bu
32	66	12.3	192	2 P94238	P94238 borrelia bu
33	66	12.3	210	2 Q44978	Q44978 borrelia bu
34	66	12.3	742	6 Q95JY6	Q95jy6 macaca fasc
35	65.5	12.2	362	17 Q57942	Q57942 pyrococcus
36	65.5	12.2	454	6 Q9XSW4	Q9xsw4 mustela vis
37	65.5	12.2	795	5 Q62006	Q62006 branchiost
38	65.5	12.2	1321	16 Q86653	Q86653 streptomyce
39	65.5	12.2	1687	11 Q35651	Q35651 mus musculu
40	65.5	12.2	1706	11 P97790	P97790 mus musculu
41	65.5	12.2	1719	11 P97789	P97789 mus musculu
42	65	12.1	158	17 Q97C56	Q97c56 thermoplasm
43	65	12.1	262	5 Q9BPS1	Q9bps1 bombyx mori
44	65	12.1	312	17 Q58122	Q58122 pyrococcus
45	65	12.1	785	10 Q9FRN0	Q9frn0 oryza sativ

ALIGNMENTS

RESULT 1
Q8TD33 PRELIMINARY; PRT; 95 AA.
ID Q8TD33
AC Q8TD33;
DC 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE Putative ligand binding protein RYD5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D.;
RT "Human RYD5, a new secretogloblin.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY026938; AAK08972.1;
SQ SEQUENCE 95 AA; 10457 MW; F629AF06C96D2392 CRC64;

Query Match 87.9%; Score 472; DB 4; Length 95;
Best Local Similarity 98.9%; Pred. No. 7.3e-45;
Matches 93; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	KGRALLLVALTLCICRMATGDNDEFFMDFLQTLVLTGTPPEELYEGLGKYNVEDAKA	60
Db	2	KGRALLLVALTLCICRMATGDNDEFFMDFLQTLVLTGTPPEELYEGLGKYNVEDAKA	61
Qy	61	AMTELKSCIDGLQPMHKAELVKLLVQLGSDGA	94
Db	62	AMTELKSCRDGLQPMHKAELVKLLVQLGSDGA	95

RESULT 2
Q05702 PRELIMINARY; PRT; 94 AA.
ID Q05702
AC Q05702;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Potential ligand-binding protein.
GN RYD5.
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FISCHER; TISSUE=OLFACTORY EPITHELIUM;
RX MEDLINE=92007724; PubMed=1915264;
RA Dear T.N., Boehm T., Keverne E.B., Rabbitts T.H.;
RT "Novel genes for potential ligand-binding proteins in subregions of
RT the olfactory mucosa";
RL EMBO J. 10:2813-2819(1991).
DR EMBL: X60661; CAA43068.1; -;
DR InterPro: IPR000329; Uterogloblin.
DR SMART: SM00096; UTG; 1.
SQ SEQUENCE 94 AA; 10401 MW; 25A4BBB4977E247 CRC64;
Query Match 64.2%; Score 344.5; DB 11; Length 94;
Best Local Similarity 75.0%; Pred. No. 1.1e-30;
Matches 69; Conservative 11; Mismatches 11; Indels 1; Gaps 1;
Qy 1 KGSRALLVLTFCICRMATGDNDEFFMDFLQTLVLTPEELVGTGKYNVNDAKA 60
Db 2 KGSRA-LLVLTVCICGLTRAEEDNEFFMDFLQTLVLTPEELVGTGKYNVNDAKA 60
Qy 61 AMTELKSCIDGLQPMHKAELVKLLVQLGSD 92
Db 61 ALTELKSCIDELQPVHKEQLVQLVQLDAQE 92
RESULT 3
Q8VD96 PRELIMINARY; PRT; 96 AA.
AC Q8VD96;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CC10 protein precursor.
GN CC10.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Gutierrez-Sagal R., Nieto A.;
RT "Cloning and sequencing of a cDNA encoding hamster uteroglobin/clara
RT cell 10 kDa protein.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: L37041; AAL31349.1; -;
DR InterPro: IPR003628; Uteroglobn_sub.
DR InterPro: IPR000329; Uterogloblin.
DR Pfam: PF01099; Uterogloblin; 1.
DR ProDom: PD012475; Uteroglobn_sub; 1.
DR SMART: SM00096; UTG; 1.
DR PROSITE: PS00404; UTEROGLOBIN_2; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 96 AA; 10509 MW; 5EB9CDD46143389 CRC64;
Query Match 16.9%; Score 90.5; DB 11; Length 96;
Best Local Similarity 27.8%; Pred. No. 0.019;
Matches 25; Conservative 17; Mismatches 43; Indels 5; Gaps 2;
Qy 7 LLVLTFCICRMATGDNDEFFMDFLQTLVLTPEELVGTGKYNVNDAKAAMTELK 66
Db 5 ITMAVVLVSVCCSSASSDTCPGFQVLEFLFMGS-ESSYEALKFYNPGLDQDSGTQLK 63

Qy 67 SCIDGLQPMHKAELVKL----LVQVLGSD 92
Db 64 KLVDTLPQKTRMINKLSEILTLSPLCNQD 93
RESULT 4
O62571 PRELIMINARY; PRT; 478 AA.
ID O62571
AC O62571; 1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Stress-responsive protein kinase (PRKSD).
OS Suberites domuncula (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Hadromerida; Suberitidae; Suberites.
OX NCBI_TaxID=55567;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller W.E.G.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96394688; PubMed=8798342;
RA Kruse M., Gamulin V., Cethovic H., Pancer Z., Mueller I.M.,
RA Mueller W.E.G.;
RT "Molecular evolution of the Metazoan protein kinase C multigene
RT family";
RL J. Mol. Evol. 43:374-383(1996).
DR EMBL: Y13101; CAA73555.1; -;
DR HSP; P24941; LHCL.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 478 AA; 54098 MW; 6638FF5FD88D6332 CRC64;
Query Match 16.3%; Score 87.5; DB 5; Length 478;
Best Local Similarity 31.0%; Pred. No. 0.31;
Matches 27; Conservative 18; Mismatches 25; Indels 17; Gaps 4;
Qy 21 TGEONDEF---FMDFLQTLVLTPEELVGTGKYNVNDAKAAMTELKSCIDGLQPMH 76
Db 238 TLKDTDFSNFSDFSIRCLVKNPEERMSATALLQHKFSKAK-PVAVLK----- 286
Qy 77 KAEVLKLLVQLGSDGAGTDYKDDDD 103
Db 287 --ELIQDAMRILEEGSGSDEEDND 311
RESULT 5
Q9D008 PRELIMINARY; PRT; 504 AA.
ID Q9D008
AC Q9D008;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Signal recognition particle 54 kDa.
GN SRP54.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

```
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR000421; FA5H_C.  
DR InterPro: IPR003961; FN_III.  
DR Pfam: PF00754; F5_F8_type-C; 1.  
DR Pfam: PF00041; fn3; 1.  
DR SMART: SM00060; FN3; 1.  
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.  
KW Complete proteome.  
SQ SEQUENCE 2104 AA; 236010 MW; 5FE34565959EED7 CRC64;  
  
Query Match 14.2%; Score 76; DB 16; Length 2104;  
Best Local Similarity 29.0%; Pred. No. 38;  
Matches 27; Conservative 10; Mismatches 34; Indels 22; Gaps  
QY 23 EDNDEF---PMFDLQTLVGTPEELYEGTLGVNYVNEDAKAAMTELKSCIDGLQPMHKAE 79  
: : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 721 KDSDAFLYDFADITLKOLLANSAGEYEVMCNAYNNNGEKFKFVSQKF-----LE 770
```

Qy	80	LVKLLVQLGSG-----DCAGTYDKDDDD	103
		: :	
Db	771	LKIQERVLSRPEFLIGNWIEDARTMLKDSDD	803
		: :	
RESULT 7			
Q9VAB8			
ID	Q9VAB8	PRELIMINARY;	PRT; 457 AA.
AC	Q9VAB8;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	CG7928 protein (L015405P).		
GN	CG7928.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		

OX	N[CBI_raxid=7221;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY.
RX	MEDLINE=20196006; PubMed=107311132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Stutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Ratton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Arrill J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Barkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liang D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,


```
DR SMART; SMO0382; AAA; 1.
DR PROSITE; PS00300; SRP54; UNKNOWN_1.
SQ SEQUENCE 500 AA; 55492 MW; C957B6F76C6A224B CRC64;

Query Match 13.58; Score 72.5; DB 11; Length 500;
Best Local Similarity 32.48; Pred. No. 15;
Matches 36; Conservative 14; Mismatches 42; Indels 19; Gaps 6;

Qy 1 KGSRALLVALLTFCICRMATGENDDEF---FMDFLQTLV-VGTPPELYEGTLGKYNNV 55
   || || || || || || || || || || || || || || || || || || || ||
Db 255 KGGGALSAAVATKSPFIIGTGEHIDNEPEKTOPFFISKILGMDIEGLID-----KVN 308
   || || || || || || || || || || || || || || || || || || || ||

Qy 56 E-----DAKAAATELKSCIDGLQPMHK--AELVKL--LVQVLGSDQAGTDY 98
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 ELKLDNDKALIEKWKHGQFTLRDMYEQFQNTMKMGPPFSQILGMIPGFGTDF 359
   || || || || || || || || || || || || || || || || || || || ||

RESULT 11
Q8QXL2 PRELIMINARY; PRT; 3071 AA.
ID Q8QXL2
AC Q8QXL2;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Polyprotein.
OS Sorghum mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=32619;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=XIAOSHAN;
RA Chen J., Chen J.P., Adams M.J.;
RT "Characterisation of potyviruses from sugarcane and maize in China.";
RL Arch. Virol. 147:0-0(2002).
DR EMBL; AJ310197; CAC84437.1;
FT CHAIN 2 233 P1 PROTEIN.
   HC-PRO PROTEIN.
FT CHAIN 234 693 P3 PROTEIN.
   P3 PROTEIN.
FT CHAIN 594 1040 P1 PROTEIN.
   HC-PRO PROTEIN.
FT CHAIN 1041 1107 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1108 1745 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1746 1798 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1799 1987 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1988 2229 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 2230 2750 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 2751 3071 P1 PROTEIN.
   P1 PROTEIN.
SQ SEQUENCE 3071 AA; 349970 MW; 2FD59A0668DE2FB7 CRC64;

Query Match 13.18; Score 70.5; DB 12; Length 3071;
Best Local Similarity 33.38; Pred. No. 2.5e+02;
Matches 22; Conservative 10; Mismatches 17; Indels 17; Gaps 3;

Qy 21 TCEDNDEEFMDFLQV-----LLVGTPEELYEGTLGKYNNVEDAKAAATELKSCIDGLQPMH 76
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2363 TGKKKD-YFMDFSQDTEKILQSCERLYEGQLGVWNGSLKAE-----IRPIE 2409
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 77 KAEVVK 82
   | | | |
Db 2410 KTEANK 2415
   | | | |

RESULT 12
Q8QXL1 PRELIMINARY; PRT; 3071 AA.
ID Q8QXL1
AC Q8QXL1;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Polyprotein.
OS Sorghum mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=32619;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
   Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
   Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblyum T.V.,
   Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBa0026L12 genomic sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC068924; AAG13498.1;
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 643 AA; 70706 MW; C695CA5DF1A36463 CRC64;

Query Match 12.98; Score 69.5; DB 10; Length 643;
Best Local Similarity 22.48; Pred. No. 44;
Matches 33; Conservative 19; Mismatches 44; Indels 51; Gaps 6;

Qy 2 GSRALLLVALLTFCICRMATGEDNDEFMDFLQTLVLTGTP---EELYE-----GTLGKY 52
   || || || || || || || || || || || || || || || || || || || ||
Db 296 GNRALTDBELMLCSEFLVAGTDSAAQLQWIMAEILVRNPISQSLYERIKSKTGGGGH 355
   || || || || || || || || || || || || || || || || || || || ||

Qy 53 NVNEDAKAAATELKSCI-DGLQ---PMH-----KAEVVKLLVQVLG 89
   || || || || || || || || || || || || || || || || || || || ||
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YUHAN;
RA Chen J., Chen J.P., Adams M.J.;
RT "Characterisation of potyviruses from sugarcane and maize in China.";
RL Arch. Virol. 147:0-0(2002).
DR EMBL; AJ310198; CAC84438.1;
FT CHAIN 2 233 P1 PROTEIN.
   HC-PRO PROTEIN.
FT CHAIN 234 693 P3 PROTEIN.
   P3 PROTEIN.
FT CHAIN 594 1040 P1 PROTEIN.
   HC-PRO PROTEIN.
FT CHAIN 1041 1107 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1108 1745 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1746 1798 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1799 1987 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 1988 2229 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 2230 2750 P1 PROTEIN.
   P1 PROTEIN.
FT CHAIN 2751 3071 P1 PROTEIN.
   P1 PROTEIN.
SQ SEQUENCE 3071 AA; 349537 MW; E3D326DE1B9ACB05 CRC64;

Query Match 13.18; Score 70.5; DB 12; Length 3071;
Best Local Similarity 33.38; Pred. No. 2.5e+02;
Matches 22; Conservative 10; Mismatches 17; Indels 17; Gaps 3;

Qy 21 TCEDNDEEFMDFLQV-----LLVGTPEELYEGTLGKYNNVEDAKAAATELKSCIDGLQPMH 76
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2363 TGKKKD-YFMDFSQDTEKILQSCERLYEGQLGVWNGSLKAE-----IRPIE 2409
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 77 KAEVVK 82
   | | | |
Db 2410 KTEANK 2415
   | | | |

RESULT 13
Q9FW93 PRELIMINARY; PRT; 643 AA.
ID Q9FW93
AC Q9FW93;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Putative cytochrome P450.
GN OSJNBa0026L12.20.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,
   Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
   Bowman C.L., Craven B., Uterback T.R., Khalak H., Feldblyum T.V.,
   Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 10 BAC OSJNBa0026L12 genomic sequence.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AC068924; AAG13498.1;
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 643 AA; 70706 MW; C695CA5DF1A36463 CRC64;

Query Match 12.98; Score 69.5; DB 10; Length 643;
Best Local Similarity 22.48; Pred. No. 44;
Matches 33; Conservative 19; Mismatches 44; Indels 51; Gaps 6;

Qy 2 GSRALLLVALLTFCICRMATGEDNDEFMDFLQTLVLTGTP---EELYE-----GTLGKY 52
   || || || || || || || || || || || || || || || || || || || ||
Db 296 GNRALTDBELMLCSEFLVAGTDSAAQLQWIMAEILVRNPISQSLYERIKSKTGGGGH 355
   || || || || || || || || || || || || || || || || || || || ||

Qy 53 NVNEDAKAAATELKSCI-DGLQ---PMH-----KAEVVKLLVQVLG 89
   || || || || || || || || || || || || || || || || || || || ||
```

```
Db 356 EVSEEDVHDPYLKAVLEGLRKHPPAHMLLPHKAAEDMDVGYLIPKGTIVNFWAENG 415
QY 90 SQ-----DAGTGD 97
Db 416 RDEKEWEPMEFPERFLPGDGEGVD 442
RESULT 14
OS1827 PRELIMINARY; PRT: 2458 AA.
AC 051827;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyketide synthase type I.
GN PLTB.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PF-5;
RX MEDLINE=98094250; PubMed=9434161;
RA Nowak-Thompson B., Gould S.J., Loper J.E.;
RT "Identification and sequence analysis of the genes encoding a
RT polyketide synthase required for pyoluteorin biosynthesis in
RT Pseudomonas fluorescens Pf-5."
RL Gene 204:17-24(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PF-5;
RX MEDLINE=99194726; PubMed=10094695;
RA Nowak-Thompson B., Chaney N., Wing J.S., Gould S.J., Loper J.E.;
RT "Characterization of the pyoluteorin biosynthetic gene cluster of
RT Pseudomonas fluorescens Pf-5."
RL J. Bacteriol. 181:2166-2174(1999).
DR EMBL; AF081920; AAC38075.1;
DR InterPro; IPR001227; Ac transferase.
DR InterPro; IPR000794; Ketoacyl-synt.
DR InterPro; IPR003880; Pantane_attach.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF0109; ketoacyl-synt; 2.
DR Pfam; PF02801; ketoacyl-synt_C; 2.
DR Pfam; PF00550; pp-binding; 2.
DR PROSITE; PS00075; ACP_DOMAIN; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
DR PROSITE; PS00098; THIOLASE_1; UNKNOWN_1.
KW Phosphopantetheine; Transferase.
SQ SEQUENCE 2458 AA; 262676 MW; AE756080AEIA5FB1 CRC64;

Query Match 12.9%; Score 69.5; DB 2; Length 2458;
Best Local Similarity 32.1%; Pred. No. 2.4e+02;
Matches 17; Conservative 14; Mismatches 19; Indels 3; Gaps 2;

QY 43 ELYEGTGLKYNVNDKAAATELKCIDGLQPMHKAELVLLVQLGSDGAG 95
Db 2239 QLLHLTQSELEQDA-AARTAMQG--SGLOPLQRSQIVQAIARVLGGGQCG 2288

RESULT 15
Q9CFY4 PRELIMINARY; PRT: 540 AA.
AC Q9CFY4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fibronectin-binding protein.
GN YNGB OR L11327.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
```

```
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=1137471;
RA Bolotin A., Wincker P., Mauer S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006365; AAK05425.1;
KW Complete proteome.
SQ SEQUENCE 540 AA; 61285 MW; 4DAC3CCCBFCF72867 CRC64;

Query Match 12.8%; Score 69; DB 16; Length 540;
Best Local Similarity 30.6%; Pred. No. 41;
Matches 22; Conservative 9; Mismatches 37; Indels 4; Gaps 2;

QY 24 DNDEFEM--DFLOTLLVGTPEELYEGTLGKYNVNDKAAATELKCIDGLQPMHKA 79
Db 294 DNAELFROKGEILLNTELNQVNDKTSVLENYTNEPIEIALNPALSPVQNAQRYFHR 353

QY 80 LVKLLVQVLGSQ 91
Db 354 KLKQAVKELGEQ 365

Search completed: April 28, 2003, 14:43:30
Job time : 26.732 secs
```